



**THÈSE DE DOCTORAT DE  
L'UNIVERSITÉ PIERRE ET MARIE CURIE**

Spécialité

**Informatique**

École doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

**M. Adrien UGON**

Pour obtenir le grade de

**DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE**

Sujet de la thèse :

**Fusion Symbolique et Données Polysomnographiques**

soutenue le Lundi 14 Janvier 2013

devant le jury composé de :

M. Jean-Gabriel GANASCIA	Directeur de thèse
M. Pierre LÉVY	Co-Directeur de thèse
M. Jean-Yves BOIRE	Rapporteur
Mme Marie-Christine JAULENT	Rapporteur
M. Stéfan DARMONI	Examineur
M. Patrick GARDA	Examineur
Mme Carole PHILIPPE	Examineur
M. Vincent RIALLE	Examineur

---

# Résumé

Au cours des dernières décennies, la médecine a nécessité des examens de plus en plus complexes pour effectuer un diagnostic et orienter vers un traitement. Il est même courant de pratiquer plusieurs examens dans des spécialités médicales différentes afin d'étudier une pathologie par des approches multiples et ainsi mieux la connaître. Cela pose des difficultés d'interprétation car les données sont parfois hétérogènes mais surtout souvent très pointues et leur traitement requiert une expertise du domaine.

Dans ce contexte, la fusion symbolique constitue une solution possible. En effet, elle a prouvé son efficacité à traiter des problèmes sur des niveaux d'abstraction aussi bien faibles qu'élevés et à élaborer une connaissance de haut niveau.

Cette thèse démontre l'efficacité de la fusion symbolique appliquée au traitement des données polysomnographiques pour l'élaboration d'un outil de support au diagnostic du Syndrome d'Apnées du Sommeil. Pour être diagnostiqué, ce trouble du sommeil nécessite une polysomnographie. Cet examen médical consiste en l'enregistrement simultané de divers paramètres physiologiques durant toute une nuit. Son interprétation nécessite l'annotation des courbes enregistrées par une analyse visuelle effectuée par un médecin spécialiste du sommeil, ce qui est une tâche chronophage et fastidieuse dont les résultats peuvent présenter quelques divergences d'un expert à l'autre. Le recours à un outil de support au diagnostic fiable permet d'augmenter le consensus. Cette thèse développe les étapes d'élaboration d'un tel outil.

# Abstract

In recent decades, medical examinations required to diagnose and guide to treatment became more and more complex. It is even a current practice to use several examinations in different medical specialties to study a disease through multiple approaches so as to describe it more deeply.

The interpretation is difficult because the data is both heterogeneous and also very specific, with skilled domain of knowledge required to analyse it.

In this context, symbolic fusion appears to be a possible solution. Indeed, it was proved to be very effective in treating problems with low or high levels of abstraction of information to develop a high level knowledge.

This thesis demonstrates the effectiveness of symbolic fusion applied to the treatment of polysomnographic data for the development of an assisted diagnosis tool of Sleep Apnea Syndrome. Proper diagnosis of this sleep disorder requires a polysomnography. This medical examination consists of simultaneously recording of various physiological parameters during a night. Visual interpretation is tedious and time consuming and there commonly is some disagreement between scorers. The use of a reliable support-to-diagnosis tool increases the consensus. This thesis develops stages of the development of such a tool.



# Remerciements

Je tiens tout d’abord à remercier mon directeur de thèse, Monsieur Jean-Gabriel GANASCIA, professeur au LIP6, responsable de l’équipe ACASA qui m’a guidé au travers du milieu de l’intelligence artificielle et m’a aidé à m’approprier la fusion symbolique. Je lui suis reconnaissant pour son exigence, son enthousiasme et ses encouragements.

J’apporte les mêmes remerciements à mon co-directeur de thèse, Monsieur Pierre P. LÉVY, praticien hospitalier dans le Département de Santé Publique de l’Hôpital Tenon (AP-HP) et membre de l’équipe INSERM URM-S 707. Sa passion, son pragmatisme et son obsession d’être utile au patient sont des leçons qu’il m’a enseignées et qui me sont désormais chères.

Je remercie chaleureusement Madame Carole PHILIPPE, praticien hospitalier responsable de l’unité d’explorations fonctionnelles du sommeil de l’Hôpital Tenon (AP-HP) pour sa gentillesse immuable, sa disponibilité et son écoute, les explications qu’elle m’a apportées et la formidable base de données qu’elle a mise à ma disposition rendant ainsi mon travail plus concret.

Je voudrais remercier également Madame Hélène AMIEL, praticien hospitalier neurologue attachée à l’unité d’explorations fonctionnelles du sommeil de l’Hôpital Tenon (AP-HP) pour son expertise, sa curiosité et sa disponibilité. Elle m’a apporté une aide précieuse pour mieux comprendre la problématique de stadification du sommeil.

Je souhaite apporter un grand merci à Mademoiselle Xue LIU, stagiaire dans l’équipe ACASA au LIP6, pour son sérieux et son efficacité. Elle a contribué à une partie des résultats exposés. Travailler avec elle fut un plaisir.

J’aimerais également remercier les membres de l’équipe ACASA au LIP6 : Madame Claire LAUDY, Monsieur Bassel HABIB, Monsieur Karan FOULADI, Monsieur Christophe JOUIS, Monsieur David PERGAMENT, Monsieur Alexandre BAZIN, Monsieur Amyn BENNAMANE, Monsieur Bin YANG et les autres stagiaires croisés au cours de ces années. Nos discussions et nos échanges furent particulièrement intéressants.

J’adresse un merci particulier à Madame Véronique LECLERC, infirmière hémovigilante dans le Département de Santé Publique, et à Madame Sabine GUESSANT, Pharmacien des Hôpitaux à la Pharmacie à Usage Intérieur, toutes les deux à l’Hôpital Tenon, pour leur soutien et leur amitié durant ces années.

Je tiens à remercier le personnel enseignant avec lequel j’ai travaillé en tant qu’ATER à l’Université Paris Ouest Nanterre La Défense (Paris X). Je pense en particulier à

Madame Claire HANEN, Madame Juliette ARNAL, Monsieur François DELBOT, Madame Bernadette DUFRÊNE, Madame Sana BEN HAMIDA, Monsieur LAURENT MESNAGER et Monsieur Bertrand LE CUN. Cette expérience m’a donné goût à l’enseignement en milieu universitaire. Leur accompagnement y a contribué.

I also would like to thank very much Mister Andrew STRANIERI, Researcher at Ballarat University in Australia, and all the Health Informatics Unit for welcoming me and organizing special meetings about the Sleep Project when I was there, and for continuing working on it after I left. I really appreciated their cheerfulness and their enthusiasm for this project.

Je tiens évidemment à remercier ma famille, à commencer par mes parents, Muriel et Pierre Yves, ma sœur Aurélie et mon frère Julien et sa femme Nadya, ainsi que mon oncle « *tonton Michel* » (UGON), pour leur présence, leur soutien, leur écoute et leurs relectures finales. Ils me sont chers et précieux en chaque instant. Sans eux, je ne serais jamais arrivé au bout.

Enfin, je remercie mes amis qui ont toujours été présents pour m’apporter l’évasion qui a été souvent nécessaire. Merci donc en premier lieu à Harold et Cédric mais aussi à Pauline, Ezilda, Steeve, Allan et Gilles. Chacun saura se souvenir d’un épisode (voire plusieurs) où leur amitié m’a permis de retrouver l’énergie indispensable pour mener à bien des travaux de thèse.

# Table des matières

Résumé . . . . .	i
Abstract . . . . .	ii
Remerciements . . . . .	iii
Table des matières . . . . .	v
<b>Introduction</b>	<b>1</b>
1 Problématique . . . . .	1
2 Contexte . . . . .	2
3 Plan . . . . .	3
<b>1 Le Syndrome d'Apnées du Sommeil</b>	<b>5</b>
1 Fusion symbolique . . . . .	5
1.1 Fusion symbolique vs Fusion numérique . . . . .	5
1.2 Stratégies de fusion . . . . .	6
1.3 Architecture du système de fusion . . . . .	6
1.4 Représentation des données . . . . .	7
2 Le Syndrome d'Apnées du Sommeil . . . . .	8
2.1 Les données polysomnographiques . . . . .	8
2.2 Interprétation des données polysomnographiques . . . . .	9
2.3 Diagnostic . . . . .	23
2.4 Traitement . . . . .	26
3 Variabilité inter- & intra-experts . . . . .	26
4 Pertinence . . . . .	28
5 Conclusion . . . . .	29
<b>2 Reconnaissance des stades de sommeil</b>	<b>31</b>
1 Données sources . . . . .	31
2 Arbre binaire de décision . . . . .	34
2.1 État de l'art . . . . .	34
2.2 Stadification automatique par arbre binaire de décision . . . . .	36
2.3 Hypothèses de travail . . . . .	37
2.4 Choix des attributs d'entrée . . . . .	41
2.5 Élaboration de l'arbre . . . . .	43
2.6 Résultats . . . . .	44
2.7 Discussion . . . . .	50
3 Fusion Symbolique . . . . .	52
3.1 Extraction de paramètres symboliques . . . . .	53
3.2 Identification des stades de sommeil par Fusion Symbolique . . . . .	72
3.3 Lissage . . . . .	79
3.4 Résultats . . . . .	79

3.5	Discussion . . . . .	83
3.6	Conclusion . . . . .	83
4	Analyse probabiliste . . . . .	85
4.1	Méthode . . . . .	85
4.2	Résultats . . . . .	86
4.3	Discussion . . . . .	86
5	Conclusion . . . . .	86
<b>3</b>	<b>Reconnaissance d'événements physiologiques</b>	<b>89</b>
1	Désaturations . . . . .	89
1.1	Description de l'algorithme . . . . .	89
1.2	Résultats . . . . .	89
1.3	Discussion . . . . .	90
2	Événements respiratoires . . . . .	91
2.1	Reconnaissance des événements ventilatoires selon leur nature . . .	91
2.2	Détection du type d'événement respiratoire par Fusion Symbolique	99
2.3	Résultats . . . . .	103
2.4	Discussion . . . . .	112
3	Aspects collaboratifs . . . . .	119
3.1	Identification automatique de Stades de Sommeil . . . . .	119
3.2	Détection des complexes K par optimisation non lisse . . . . .	125
3.3	Détection des complexes K utilisant un apprentissage hybride-synergique . . . . .	130
3.4	Conclusion . . . . .	139
4	Conclusion . . . . .	140
<b>4</b>	<b>Des caractéristiques à la décision</b>	<b>141</b>
1	Bibliographie . . . . .	141
2	Données initiales . . . . .	145
2.1	Données contextuelles : Stades de sommeil et Position . . . . .	146
2.2	Données événementielles : Apnées, désaturations et micro-éveils . .	146
3	Méthode . . . . .	148
4	Résultats . . . . .	148
5	Discussion . . . . .	160
6	Conclusion . . . . .	162
<b>5</b>	<b>Visualisation</b>	<b>163</b>
1	OptiSAS . . . . .	164
1.1	Méthode . . . . .	164
1.2	Résultats . . . . .	166
1.3	Évaluation . . . . .	167
1.4	Discussion . . . . .	168
2	SAS-Map . . . . .	168
2.1	Méthode . . . . .	169
2.2	Résultats . . . . .	171
2.3	Discussion . . . . .	172
3	Conclusion . . . . .	172
	<b>Conclusion</b>	<b>173</b>

Bibliographie	177
---------------	-----



# Introduction

## 1 Problématique

La problématique à laquelle nous nous proposons de nous confronter est celle posée par la question suivante : « Est-il pertinent d'utiliser la fusion symbolique pour des problèmes biomédicaux ? » Nous avons choisi d'y répondre en testant la fusion symbolique sur des données polysomnographiques, c'est-à-dire issues de l'examen du sommeil, dans le but de créer un outil de support au diagnostic du Syndrome d'Apnées du Sommeil.

La fusion de données est utilisée dans des contextes où plusieurs capteurs enregistrent, en parallèle, des données dans un même environnement pour observer un phénomène. Chaque capteur peut alors être considéré comme une fenêtre qui permet d'observer le phénomène selon un angle et une précision qui lui sont propres. Seule la mise en commun des informations issues de toutes les sources d'information permet de reconstituer une vue d'ensemble du phénomène.

De nombreuses approches numériques ont été expérimentées, souvent dans l'hypothèse qu'en travaillant au plus proche des signaux recueillis, alors meilleur serait le résultat, car la perte d'information utile serait minimale. Les différents capteurs sont alors souvent utilisés pour confirmer, ou infirmer, ce qui est observé sur l'une des autres sources. Nous croyons au contraire que l'approche symbolique présente des avantages : d'une part, c'est une technique qui est moins sensible au bruit et, d'autre part, l'abstraction de l'information, vers des niveaux de plus en plus élevés permet de construire la vue d'ensemble en profitant de la contribution de chaque source, mais aussi des connaissances apportées par les experts du domaine.

Le Syndrome d'Apnées du Sommeil est une pathologie qui se prête bien à cette approche. En effet, la polysomnographie est aujourd'hui l'examen médical de référence pour son diagnostic et elle consiste en le recueil simultané de nombreux paramètres physiologiques au cours du sommeil d'un individu. La durée de l'examen peut dépasser les 10 heures et le nombre de capteurs utilisés varie généralement de 15 à 30. L'interprétation des signaux nécessite, d'une part, la mise en relation des événements observés sur les différentes sources, mais aussi, d'autre part, le calcul d'indicateurs en fonction du temps, en regroupant éventuellement certaines périodes discontinues. Aujourd'hui réalisée quasiment intégralement visuellement par un expert, conformément aux recommandations, l'interprétation des données polysomnographiques est une tâche fastidieuse et chronophage. Les analyses automatiques ne sont pas considérées comme suffisamment fiables pour être utilisées en substitut, et même parfois en simple complément, d'une analyse visuelle de l'expert.

## 2 Contexte

Cette thèse a été réalisée dans un contexte pluridisciplinaire, entre plusieurs laboratoires experts dans des domaines différents. Ayant travaillé en collaboration entre trois unités, il m’a également été nécessaire de m’adapter à chaque interlocuteur selon son domaine. En effet, chaque laboratoire a ses propres méthodes, ses objectifs et ses points de vue qui lui sont spécifiques.

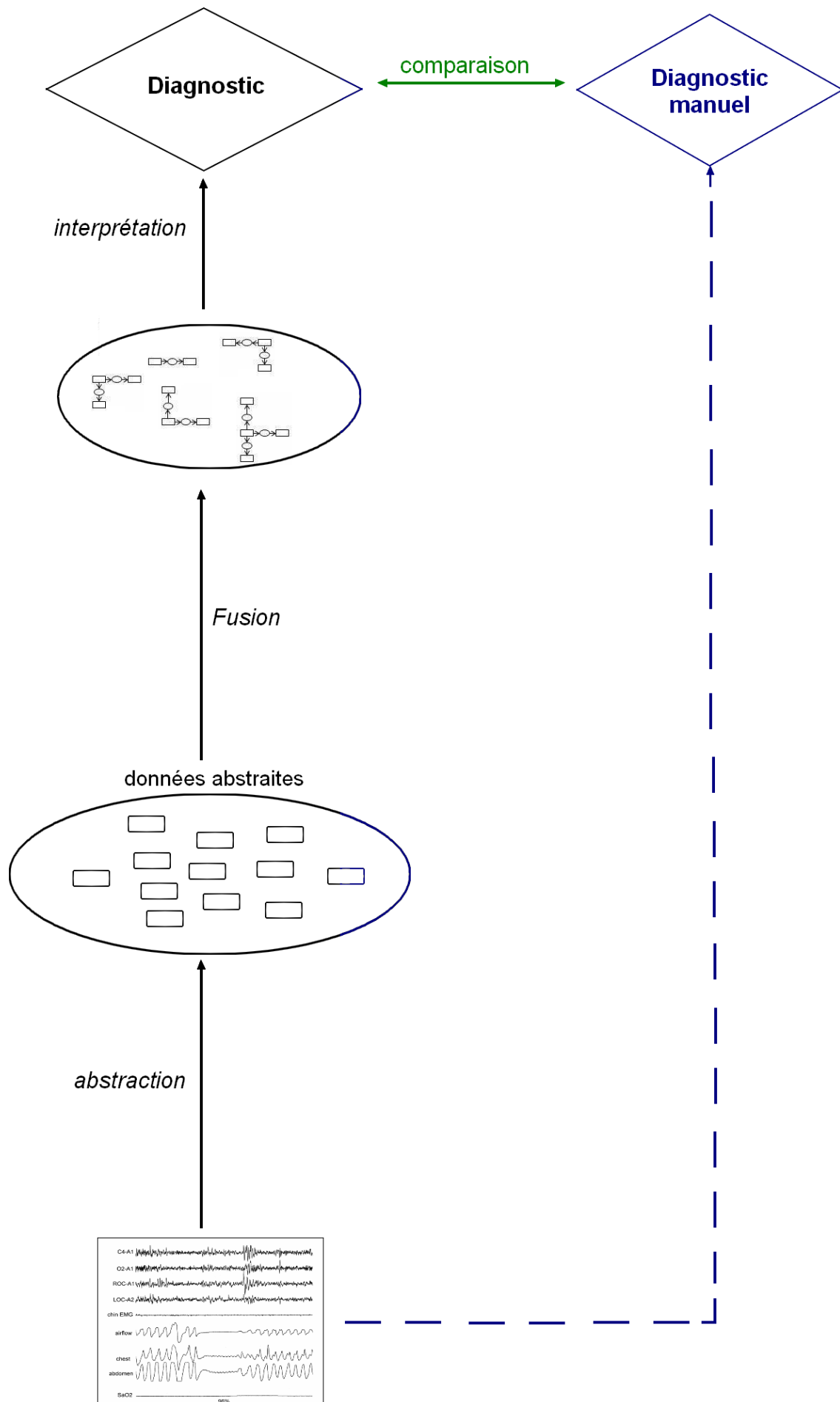
- L’équipe ACASA du LIP6, investie dans la recherche en Informatique, met en jeu des méthodes informatiques d’Intelligence Artificielle de pointe au service de projets de gestion de la connaissance. Elle a en particulier une longue expérience dans la représentation des données biologiques et médicales, dans l’analyse sémantique du langage naturel mais aussi dans la fusion de données de haut niveau. Elle a notamment apporté dans ce domaine l’utilisation de la fusion avec des graphes conceptuels. Elle travaille aussi sur l’analyse de graphes, en particulier dans le domaine des réseaux sociaux, pour en extraire de l’information. Cette équipe est dirigée par le Professeur Jean-Gabriel GANASCIA.
- Le département de Santé Publique (DSP) de l’Hôpital Tenon, Centre Hospitalier Universitaire (CHU) de l’Assistance Publique - Hôpitaux de Paris (AP-HP), lié à l’unité URM-S707 de l’INSERM, regroupe 6 médecins chercheurs, une infirmière et une chargée de mission associés dans des activités de recherche médicale, de biostatistiques, de vigilance et prévention des risques. Plusieurs experts en informatique médicale y travaillent, en particulier le Docteur Pierre LÉVY, co-encadrant de cette thèse. Le département de Santé Publique de l’Hôpital Tenon, s’occupe de recherche en informatique médicale avec un axe fort dans la visualisation de l’information médicale. Ses objectifs sont de l’ordre de l’amélioration de l’activité médicale, du travail de l’équipe soignante à la qualité de vie du patient. Pour cela, les champs à optimiser sont ceux de la précision et la rapidité du diagnostic, de la communication entre les différents acteurs ou du confort du patient.
- Le service d’explorations fonctionnelles du Sommeil de l’Hôpital Tenon, dirigé par le Docteur Carole PHILIPPE est avant tout un service clinique dont l’objectif est une prise en charge réussie du maximum de patients atteints de troubles du sommeil, principalement de Syndrome d’Apnées du Sommeil, conformément aux règles reconnues par la communauté internationale de médecine du sommeil. Il reçoit 2000 patients par an, la moitié en consultations et l’autre moitié pour des examens. Il est constitué d’un praticien hospitalier, de 5 praticiens attachés (équivalents de 9 demi-journées par semaine), de deux infirmières, d’une aide-soignante (partagée avec le service d’explorations fonctionnelles) et d’une secrétaire à mi-temps (partagée également avec l’ensemble du service). L’unité est membre du réseau « Hypnos Ville-Hôpital ».

Ces trois équipes partagent des points communs : d’une part, la passion pour leur domaine et d’autre part, des contraintes en terme de moyens, financiers et humains, à gérer. S’il est évident que la mise en relation de ces trois univers crée des difficultés de compréhension, elle est avant tout une richesse et il est à noter que chacun a su faire des efforts pour assurer la bonne cohésion du projet.



### 3 Plan

Dans un premier temps, nous allons présenter la fusion symbolique et le contexte médical du Syndrome d'Apnées du Sommeil. La fusion symbolique présente l'intérêt de pouvoir être utilisée sur des objets symboliques, indépendamment des méthodes utilisées pour leur recueil. La deuxième partie sera consacrée à la reconnaissance automatique des stades de sommeil ; nous y présenterons notamment deux algorithmes différents. Dans le chapitre 3, nous verrons un algorithme de reconnaissance automatique des événements caractéristiques du Syndrome d'Apnées du Sommeil, en particulier les événements respiratoires et les désaturations. Le chapitre 4 est central car c'est celui qui présente le cœur de la thèse, c'est-à-dire la fusion des événements symboliques obtenus aux étapes précédentes. La fusion symbolique génère des résultats complexes qu'il faut présenter à l'expert médical dans un référentiel pertinent, seul moyen de bénéficier de toute l'information sémantique intrinsèque à ces résultats, c'est ce que nous verrons dans le chapitre 5.



# Chapitre 1

## Le Syndrome d'Apnées du Sommeil : un domaine d'application favorable à la fusion symbolique

### 1 Fusion symbolique

#### 1.1 Fusion symbolique vs Fusion numérique

Il existe deux approches de fusion automatique de données :

- la fusion de données de bas niveau qui consiste à fusionner des données à l'état brut, c'est-à-dire telles qu'elles sont recueillies ; ce sont des informations d'un bas niveau sémantique.
- la fusion symbolique, de haut niveau, qui se fait après abstraction des données recueillies. Cette fusion permet d'utiliser des données hétérogènes d'un haut niveau sémantique. La fusion symbolique permet d'obtenir une connaissance de haut niveau.

Si de nombreuses méthodes ont été proposées pour effectuer la fusion de données brutes, elles ont également apporté peu de satisfaction. Nous nous intéresserons donc davantage à la fusion symbolique. En effet, l'équipe ACASA a une longue expérience dans l'utilisation des méthodes symboliques et il m'a été offert d'en bénéficier tout au long de la durée de ce doctorat. Fort de ses expériences dans d'autres contextes, et au vu des résultats que j'avais obtenus en appliquant des méthodes de visualisation au cours d'un stage ingénieur sous la responsabilité du Docteur Pierre LÉVY<sup>1</sup> [74], le Professeur Jean-Gabriel GANASCIA a eu l'idée de traiter les données polysomnographiques avec des méthodes symboliques, en particulier la fusion, et m'a présenté des outils pour y procéder. Je me les suis appropriés et les ai adaptés à la problématique du Syndrome d'Apnées du Sommeil. Ce travail est aussi nourri par les apports scientifiques issus des discussions avec Claire LAUDY et Bassel HABIB, doctorants dans l'équipe et ayant utilisé chacun la fusion symbolique.

Ces méthodes que nous allons développer relevant de la fusion symbolique constitueront le cœur de la thèse.

---

1. Praticien Hospitalier au Département de Santé Publique de l'Hôpital TENON (AP-HP)

La fusion symbolique nécessite de définir, au préalable :

- un ensemble de concepts ;
- un ensemble de relations entre les concepts ;
- un formalisme de représentation de la connaissance ;
- les règles à utiliser pour la fusion.

## 1.2 Stratégies de fusion

Les travaux de Claire LAUDY portent sur la fusion d'informations, de haut niveau, en utilisant des graphes conceptuels [39], en s'intéressant plus particulièrement à la jointure maximale [38]. Elle a également montré l'importance de choisir les bonnes stratégies pour fusionner les données en introduisant des connaissances sémantiques pour l'élaboration des règles de fusion [37]. Ces règles sémantiques dépendent fortement du domaine d'application ; c'est pourquoi elle a appliqué ses travaux, dans un premier lieu, aux programmes télévisés [36] puis à l'exploitation des informations présentes dans les carnets d'expériences de Claude BERNARD, en collaboration avec Bassel HABIB [29,40].

## 1.3 Architecture du système de fusion

Dans [19], Belur V. DASARATHY fait une analyse détaillée des architectures des systèmes de fusion.

Il définit, dans un premier temps, trois niveaux de fusion : données, caractéristiques et décisions. Il fait le constat que de nombreux systèmes de fusion travaillent au niveau des données, là où il y a le plus d'informations mais aussi le plus de bruit. Ainsi, il est considéré que la perte d'informations est minimale et que le résultat de la fusion en sera meilleur, le système de fusion étant conçu généralement comme une validation par consolidation des décisions prises au niveau de chaque capteur. Dans une telle architecture, chaque capteur a alors une contribution identique. Ceci peut s'expliquer par la capacité supérieure de la machine par rapport à l'Homme à extraire des caractéristiques issues de larges volumes de données, et à la supériorité supposée de l'Homme (encore plus de l'expert), à prendre une décision comparée à la machine.

Belur V. DASARATHY préconise de reconsidérer les architectures des systèmes de fusion afin que les étapes de fusion des niveaux supérieurs (caractéristiques, décisions) aient une contribution plus importante au résultat final., tout en évitant qu'une source d'information défaillante à ce niveau fasse échouer le système. Il recommande, à cet effet, d'utiliser des sources complémentaires, c'est-à-dire des sources dont le spectre d'information se recouvre.

L'information doit ainsi se propager de proche en proche depuis les données brutes, afin de profiter du maximum de détails, vers la décision finale, prise le plus tard possible en profitant à chaque étape, de l'information provenant des différentes sources. Pour minimiser le risque d'échec du système, l'architecture doit avoir une utilisation adéquate des différents niveaux de fusion et des différentes sources d'information.

Enfin, Belur V. DASARATHY considère la fusion temporelle comme un axe orthogonal aux autres problèmes de fusion. Celle-ci peut intervenir aux trois niveaux définis précédemment (données, caractéristiques et décisions) et utilise généralement des techniques différentes.

## 1.4 Représentation des données

Nous avons choisi de représenter les données sous la forme de graphes conceptuels. Les graphes conceptuels ont été introduits par John F. SOWA en 1984 [70]. Ils sont utilisés comme formalisme de représentation des connaissances et de raisonnement. Nous utiliserons ici pour ontologie une hiérarchie des concepts et des relations. Sur la figure 1.1, nous donnons la hiérarchie des concepts que nous utilisons.

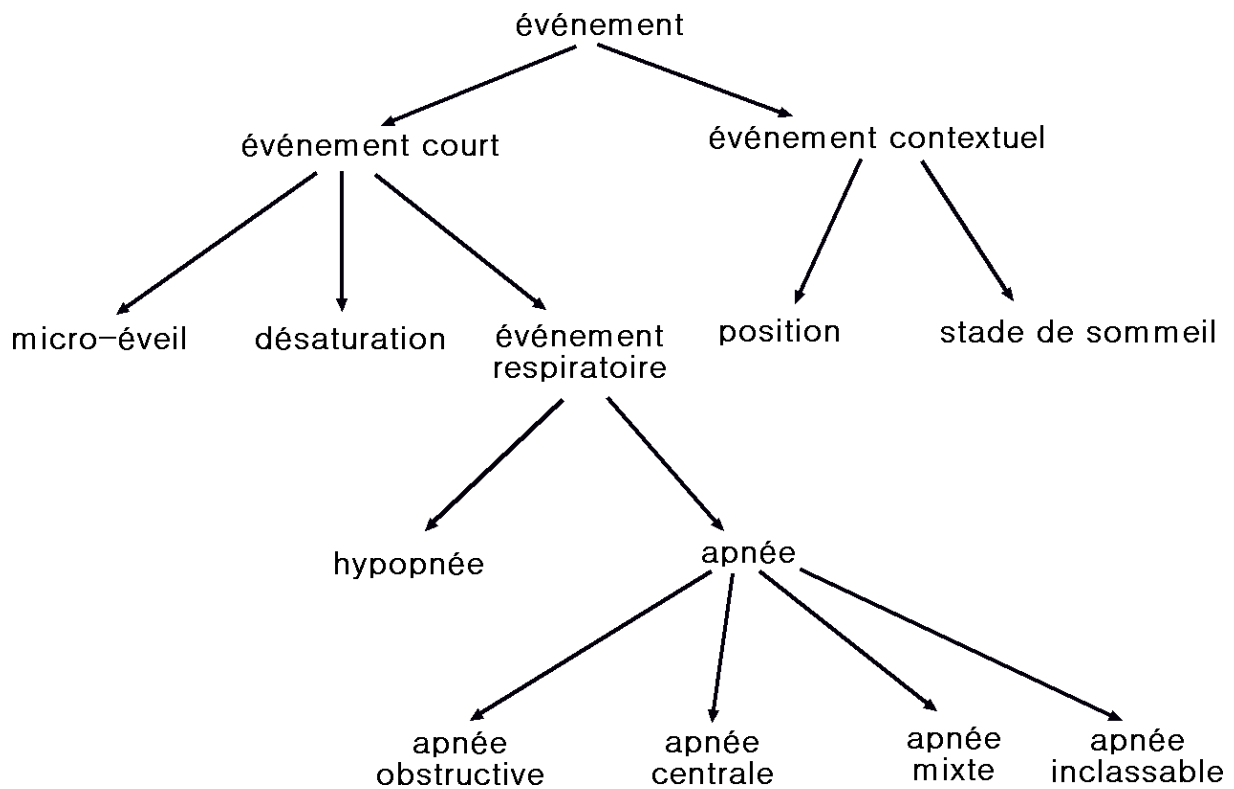


FIGURE 1.1 – Hiérarchie des concepts

## 2 Le Syndrome d'Apnées du Sommeil

### 2.1 Les données polysomnographiques

Le Syndrome d'Apnées du Sommeil (SAS) est un trouble du sommeil qui se caractérise par la survenue de cessations, totales ou partielles, du flux respiratoire au cours du sommeil. Il entraîne une fragmentation du sommeil, qui a pour conséquences une somnolence diurne excessive, des déficits cognitifs ainsi qu'une augmentation de la morbidité et la mortalité cardiovasculaires et cérébrovasculaires. Sa prévalence est de 4% chez les hommes d'âge moyen et 2% chez les femmes d'âge moyen [81].

La polysomnographie est l'examen médical standard d'exploration du sommeil. Il sert à recueillir simultanément, pendant une nuit, différents paramètres physiologiques. Il est utilisé pour le diagnostic de troubles du sommeil, comme, par exemple, le Syndrome d'Apnées du Sommeil (SAS) (cf Figure 1.2).

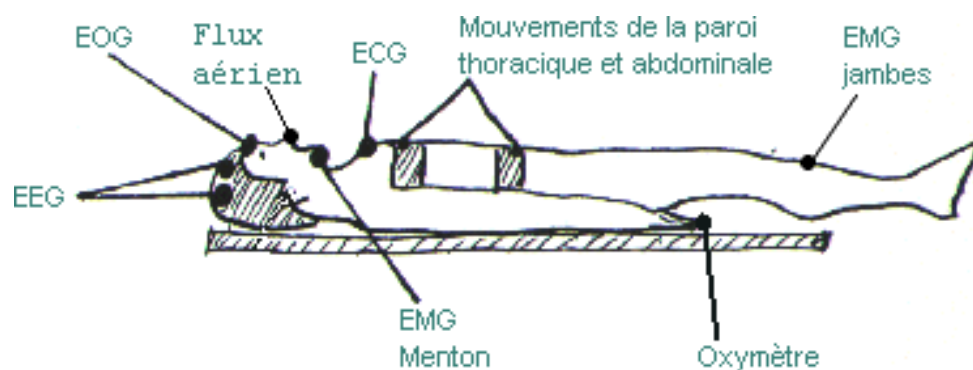


FIGURE 1.2 – Examen polysomnographique

Lors d'une polysomnographie, sont enregistrés les paramètres suivants :

- l'activité neurologique,
  - l'activité électrique cérébrale : au moins 3 voies électroencéphalographiques (EEG) ;
  - le mouvement des yeux : 2 voies électro-oculographiques (EOG) ;
  - le tonus musculaire : une voie électromyographique (EMG Menton).
- l'activité ventilatoire,
  - le flux respiratoire (canule de pression nasale, thermistance buccale ou bruits trachéaux) ;
  - les efforts respiratoires (2 ceintures thoraco-abdominales).
- l'activité cardiovasculaire,
  - l'activité électrique cardiaque (ECG) ;
  - (parfois) la photopléthysmographie de pouls.
- l'oxymétrie de pouls (ou saturométrie),
  - la saturation du sang en Oxygène ( $SpO_2$ ) ;
  - le pouls.
- les mouvements de jambes,
  - le tonus musculaire des jambes (EMG tibial antérieur).
- la position du corps à 5 positions,
  - décubitus dorsal ;
  - décubitus ventral ;
  - décubitus latéral droit ;
  - décubitus latéral gauche ;
  - debout.

## 2.2 Interprétation des données polysomnographiques

Une fois l'enregistrement terminé, le médecin doit interpréter les courbes. L'American Academy of Sleep Medicine a défini dans [31] les règles à utiliser pour l'interprétation des courbes. Cette opération s'appelle le *scoring*.

Il consiste en six lectures successives des données, pour y repérer des événements physiologiques symptomatiques du Syndrome d'Apnées du Sommeil :

1. stades de sommeil ;
2. désaturations ;
3. micro-éveils ;
4. mouvements de jambes ;
5. événements respiratoires ;
6. recherche de la nature spontanée ou liée des micro-éveils.

L'AASM recommande également le scorage des événements cardiaques :

- tachycardies ;
- bradycardies ;
- asystolies ;
- tachycardie à complexes larges ;
- tachycardie à complexes étroits ;
- fibrillation auriculaire.

### Le scorage des stades de sommeil

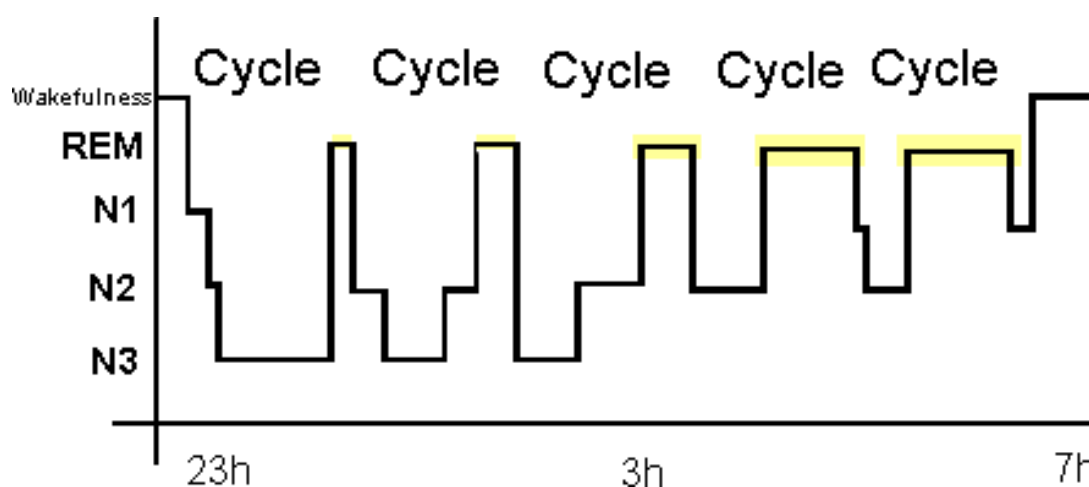


FIGURE 1.3 – Cycles et stades de sommeil

Le sommeil n'est pas uniforme. Il est normalement composé de 5 à 6 cycles successifs sur une nuit. Chaque cycle dure environ 90 minutes. Chacun de ces cycles est subdivisé en stades de sommeil, assimilables à la profondeur du sommeil (cf Figure 1.3).

Le scorage des stades de sommeil consiste à identifier, par analyse visuelle des courbes neurologiques (*EEG, EOG, EMG Menton*), le stade de sommeil d'une portion de signal. L'AASM demande de procéder à une analyse par fenêtres de 30 secondes, appelées *époques* [31].

Cinq stades de sommeil sont définis :

- l'éveil : W (*Wakefulness*),
- Le sommeil paradoxal : R (*REM*<sup>2</sup>),
- Le sommeil lent : N (*NREM*),
  - Le sommeil lent léger ;
    - N1 (*NREM 1*) ;
    - N2 (*NREM 2*).
  - Le sommeil lent profond ;
    - N3 (*NREM 3*).

---

2. REM : **R**apid **E**ye **M**ovements



Une analyse attentive de la fréquence des signaux EEG est demandée. On utilisera notamment les plages de fréquences du tableau 1.1.

TABEAU 1.1 – Plages de fréquences EEG

Plage de fréquences	label
0-4 Hz	$\delta$
4-7 Hz	$\theta$
8-12 Hz	$\alpha$
12-30 Hz	$\beta$

## Stade W

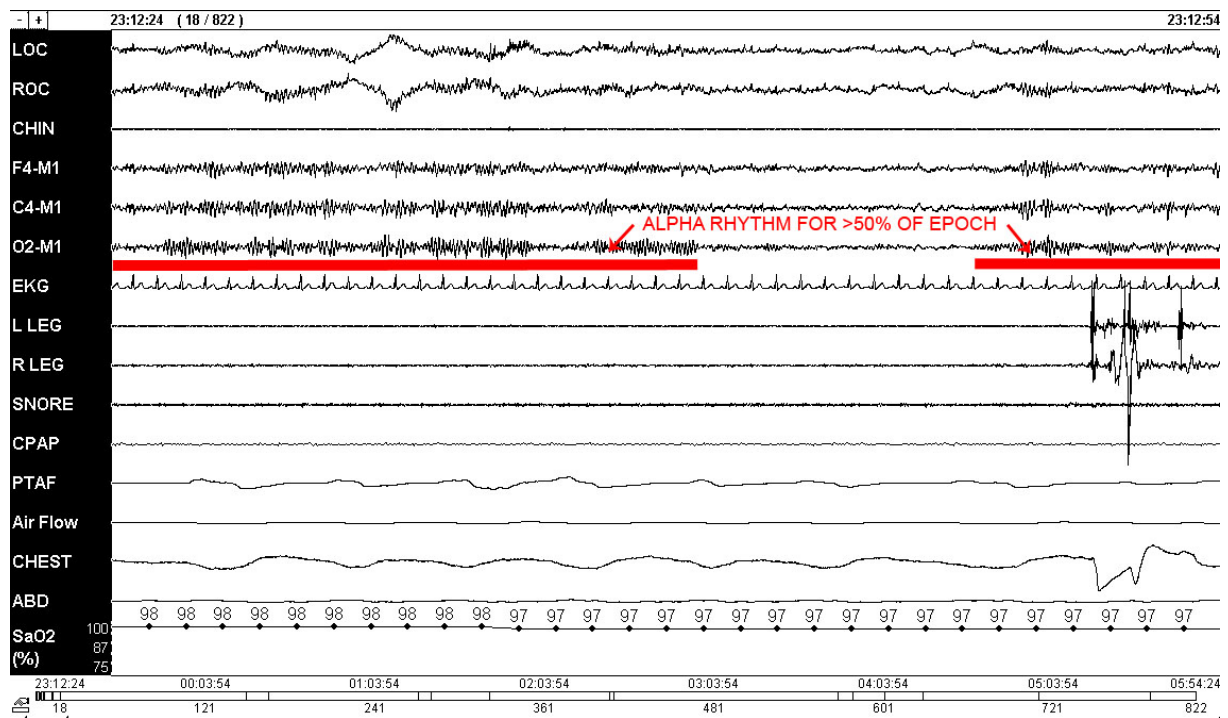


FIGURE 1.4 – Éveil (Stade W)

Le stade W correspond à l'éveil. Il est caractérisé par :

- la présence du rythme alpha sur plus de 50% de l'époque,
- en cas d'absence de rythme alpha clairement identifiable,
  - des clignements des yeux à la fréquence de 0,5-2Hz ;
  - des *Mouvements de lecture des yeux*<sup>3</sup> ;
  - l'association irrégulière de *mouvements oculaires rapides*<sup>4</sup> avec un tonus musculaire normal ou élevé.

3. Trains de signaux constitué d'une phase lente suivie d'une phase rapide dans la direction inverse

4. Mouvements oculaires rapides, conjugués, irréguliers et fortement pointus avec une déflexion initiale d'une durée inférieure à 500ms

## Stade R

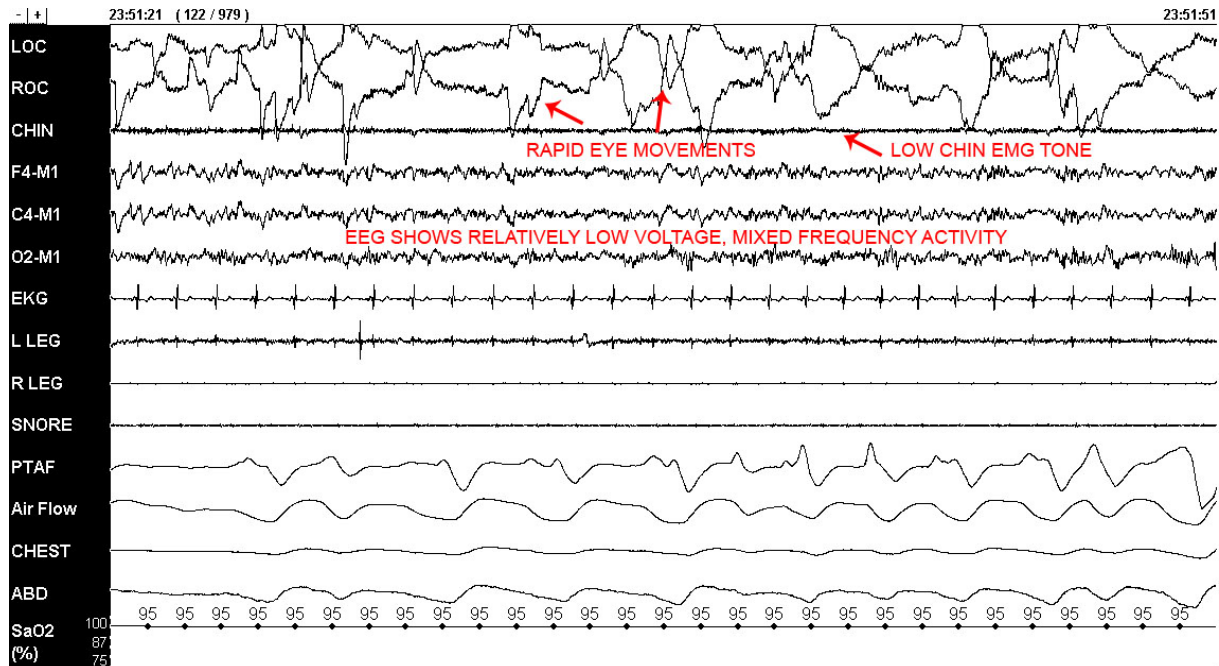


FIGURE 1.5 – Sommeil Paradoxal (Stade R)

Le stade R correspond au sommeil paradoxal. Il représente 20% du Temps Total de Sommeil. On l'observe davantage en fin de nuit.

Il est caractérisé par :

- des signaux EEG de faible amplitude avec un grand mélange de fréquences ;
- un faible Tonus musculaire au menton ;
- des mouvements oculaires rapides.

## Stade N1

Le stade N1 est le premier des 2 stades du sommeil lent léger. Il représente 5% du Temps Total de Sommeil. Il est caractérisé par :

- chez les sujets générant du rythme alpha, le remplacement du rythme alpha par des signaux EEG de faible amplitude composés de différentes fréquences sur plus de 50% de l'époque ;
- chez les sujets ne générant pas de rythme alpha,
  - des ondes rapides à une fréquence de 4-7Hz (ondes  $\theta$ ) avec ralentissement des ondes lentes d'au moins 1Hz ;
  - la présence de *pointes vertex*<sup>5</sup> ;
  - Mouvements oculaires lents.

5. Une *onde vertex* est une onde pointue aux contours nets d'une durée de 0,5 seconde

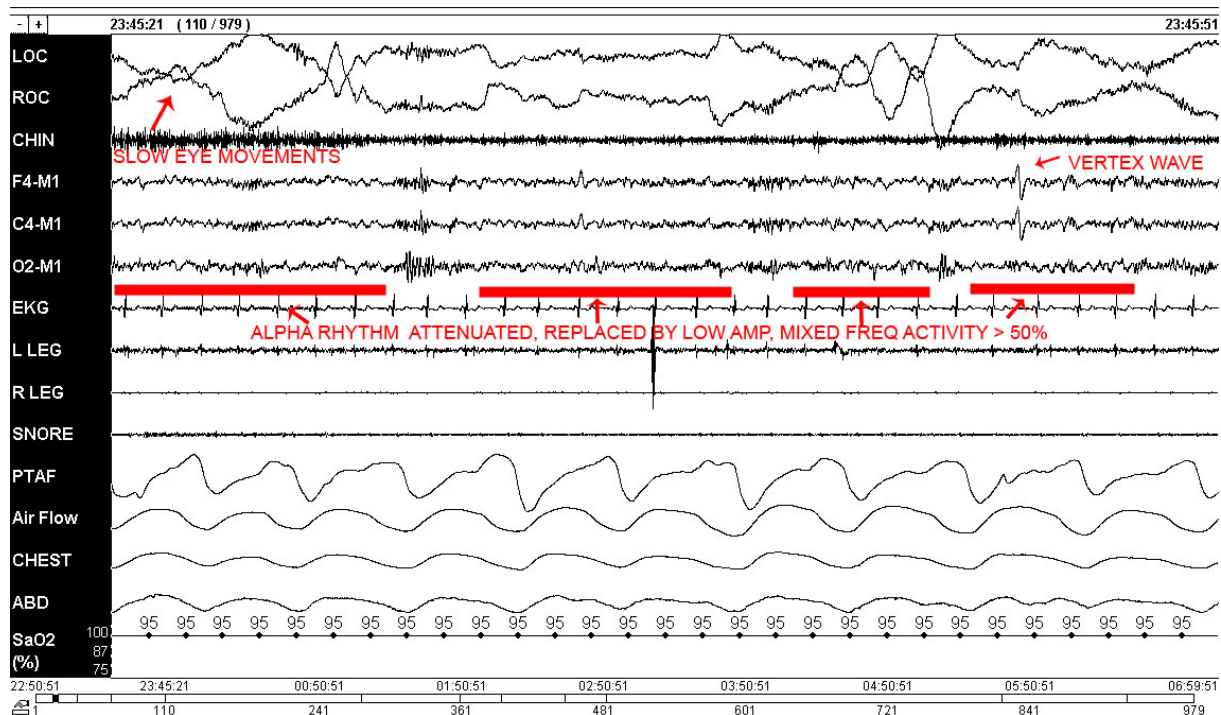


FIGURE 1.6 – Stade N1

## Stade N2

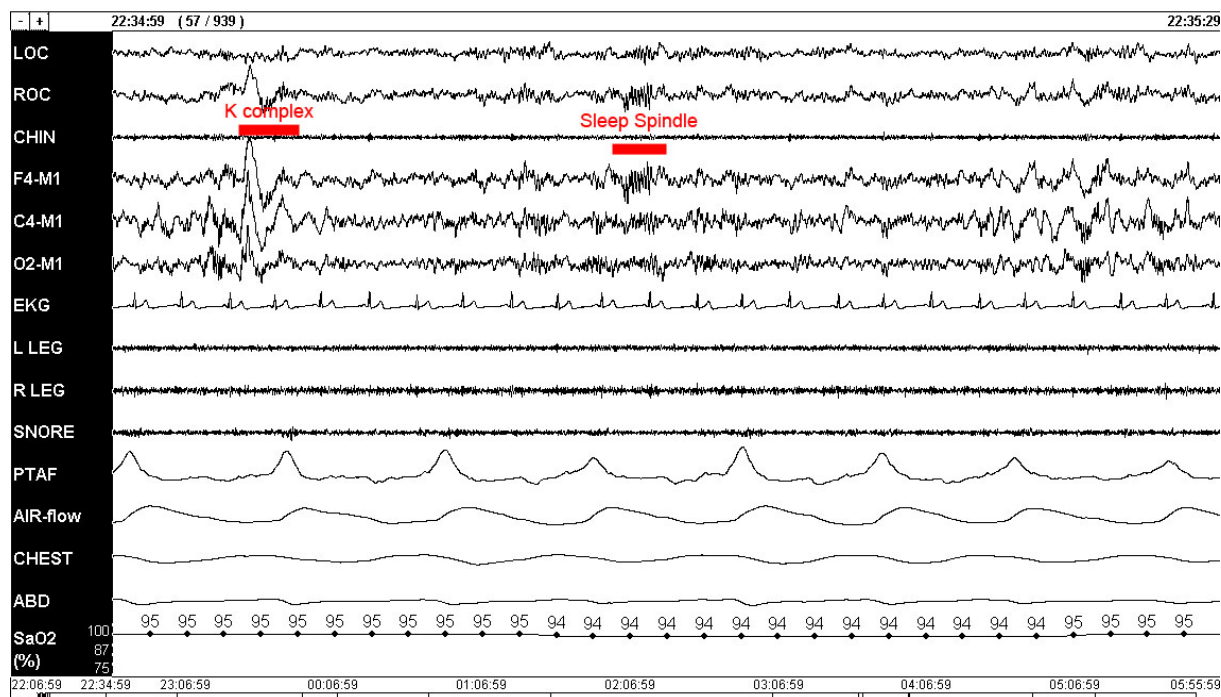


FIGURE 1.7 – Stade N2

Le stade N2 est le second des 2 stades du sommeil lent léger. Il représente 50% du Temps de Sommeil Total. Il est caractérisé par :

- la présence d'un ou plusieurs *complexes K*<sup>6</sup> (*K complex*= non-associés à un micro-éveil ;
- la présence d'un ou plusieurs *fuseaux du sommeil*<sup>7</sup> (*sleep-spindles*).

### Stade N3

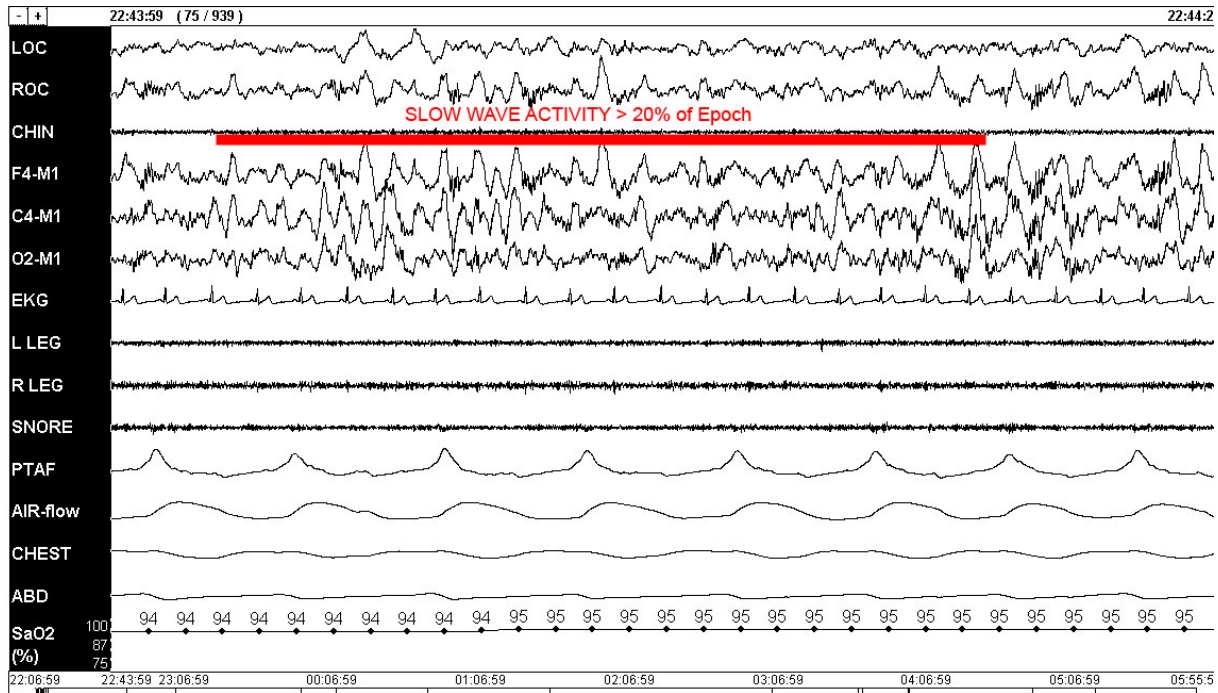


FIGURE 1.8 – Stade N3

Le stade N3 est le sommeil lent profond. Il représente 25% du Temps de Sommeil Total.

Il est caractérisé par :

- la présence d'au moins 20% d'*ondes lentes*<sup>8</sup>.

### Désaturations

La saturation du sang en Oxygène représente le taux d'Oxygène dans le sang. Idéalement, il est proche de 100%. En cas d'apnée, il diminue, ce qui peut affecter les organes vitaux. Ceci explique l'augmentation de la morbidité et la mortalité cardiovasculaire [48] et cérébrovasculaire [6] chez les patients atteints d'un Syndrome d'Apnées du Sommeil. Dans le cadre du Syndrome d'Apnées du Sommeil, une désaturation est significative si la baisse est d'au moins 3%.

6. un *complexe K* est défini comme une onde négative pointue bien délimitée immédiatement suivie par une composante positive se détachant sur le fond EEG, avec une durée totale de plus de 0,5 seconde

7. un *fuseau du sommeil* est défini comme un train d'ondes distinctes d'une fréquence de 11 à 16Hz d'une durée de plus de 0,5 seconde

8. Les *ondes lentes* sont définies comme des ondes d'une fréquence de 0,5 à 2 Hz avec une amplitude pic-à-pic d'au moins 75µV.

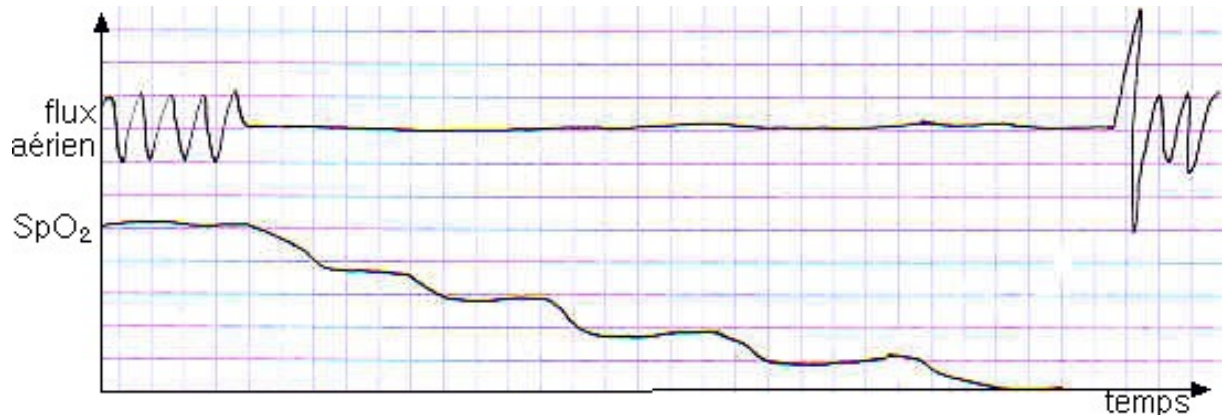


FIGURE 1.9 – Désaturation

### Micro-éveils

Les micro-éveils sont assimilables à de courts éveils avec une augmentation de l'activité cérébrale. Le dormeur n'est pas conscient de leur survenue.

Ils fragmentent le sommeil et peuvent influencer considérablement la structure du sommeil. En effet, à cause d'un sommeil trop fragmenté, certains patients ont une prédominance de sommeil lent léger (N1 et N2) avec disparition de sommeil lent et paradoxal.

Cette fragmentation du sommeil entraîne une somnolence diurne excessive, des déficits cognitifs [50] et accroît le risque d'accident de la route [25].

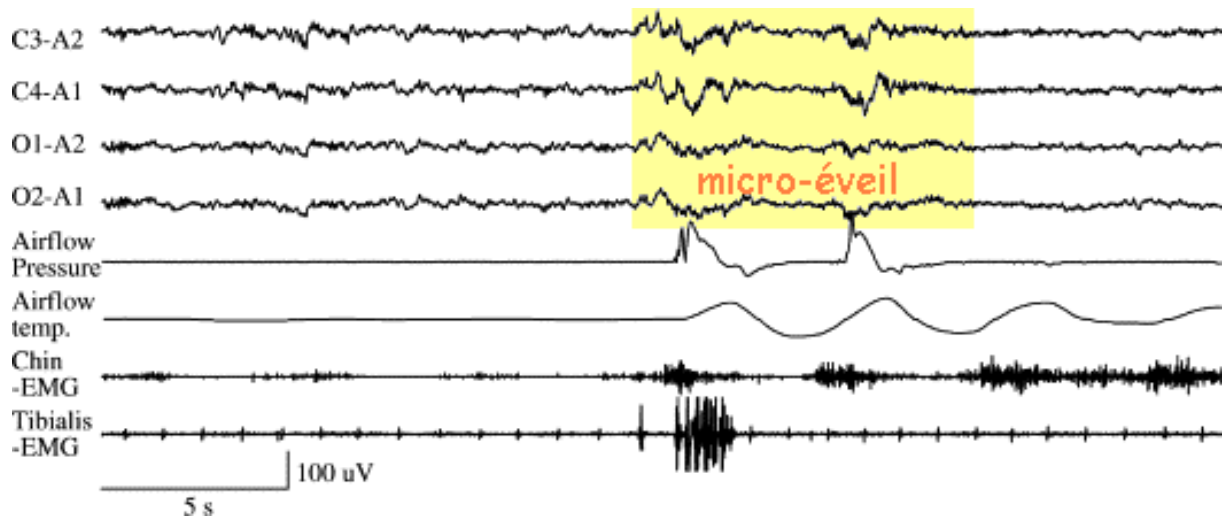


FIGURE 1.10 – Micro-éveil

D'après [31], les micro-éveils sont définis comme des événements survenant dans les stades N1, N2, N3 ou R après au moins 10 secondes de sommeil stable. Il s'agit d'un changement brusque de la fréquence de l'EEG, incluant des ondes alpha, thêta et/ou des ondes de fréquences supérieures à 16 Hz (mais pas de fuseaux), pour une durée entre 3 et 15 secondes.



En sommeil paradoxal (stade R), il est nécessaire d'observer une augmentation du tonus musculaire au menton pendant au moins une seconde.

### Mouvements de jambes

#### Mouvements de jambes

D'après [31], les mouvements de jambes sont définis par l'ensemble des critères suivants :

1. La durée minimale d'un mouvement de jambe est 0,5 seconde.
2. La durée maximale d'un mouvement de jambe est 10 secondes.
3. L'amplitude du signal EMG doit avoir une augmentation minimale de  $8\mu V$ .
4. Un mouvement de jambe commence à l'instant où il y a une augmentation de  $8\mu V$  par rapport au reste du signal l'EMG.
5. Un mouvement de jambe se termine au début d'une période d'au moins 0,5 seconde pendant laquelle l'EMG ne dépasse pas de plus de  $2\mu V$  le reste du signal EMG.

#### Mouvements périodiques de jambes

D'après [31], les mouvements périodiques de jambes (MPJ) sont définis par l'ensemble des critères suivants :

1. Le nombre minimum de mouvements de jambes consécutifs pour définir une série de Mouvements Périodiques de Jambes est 4.
2. La durée minimale entre deux mouvements de jambes consécutifs<sup>9</sup> pour qu'ils soient pris en compte dans une série de Mouvements Périodiques de Jambes est 5 secondes.
3. La durée maximale entre deux mouvements de jambes consécutifs<sup>9</sup> pour qu'ils soient pris en compte dans une série de Mouvements Périodiques de Jambes est 90 secondes.
4. Lorsque des mouvements de jambes enregistrés sur 2 jambes différentes commencent à moins de 5 secondes d'intervalle, on compte un seul mouvement de jambe.

Nous n'utiliserons pas les mouvements de jambes dans nos travaux.

### Événements respiratoires

#### Apnées

D'après [31], la détection des apnées se fait sur un signal recueilli par une pression nasale.

Une apnée est définie par l'ensemble des critères suivants :

1. cessation de flux respiratoire durant au moins 10 secondes ;
2. baisse d'amplitude du signal d'au moins 90% pendant au moins 90% de la durée de l'apnée.

La Figure 1.12 représente un schéma de courbes polysomnographiques avec la mise en évidence d'une apnée en rouge (suivie d'un micro-éveil en vert et d'une désaturation, en bleu).

---

9. La durée entre deux mouvements de jambes consécutifs est définie comme la durée entre les débuts des deux mouvements de jambes

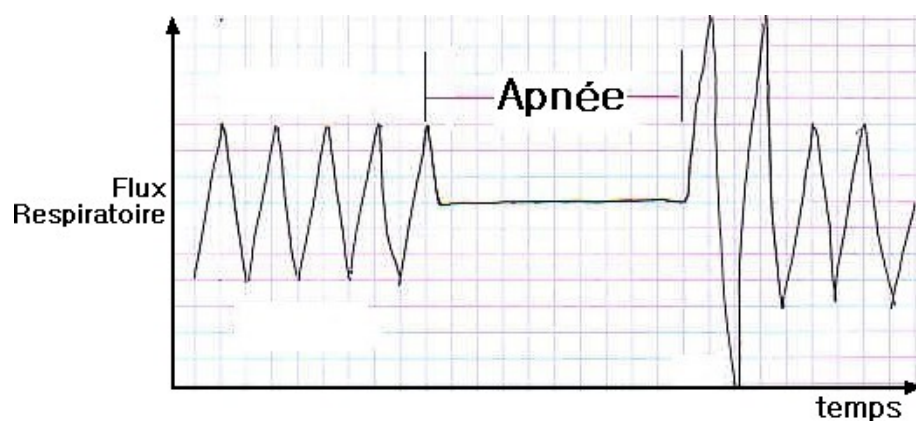


FIGURE 1.11 – Apnée

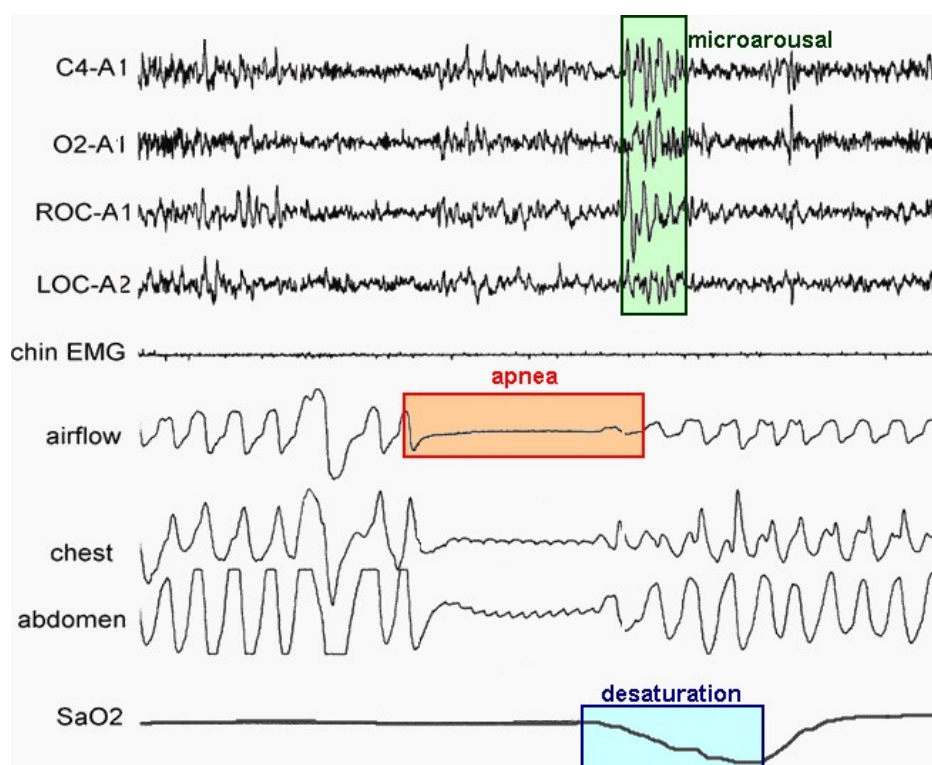


FIGURE 1.12 – Courbes polysomnographiques [schéma]

La Figure 1.13 est un aperçu d'écran de courbes polysomnographiques réelles. On peut y voir une apnée, mise en évidence par un fond vert.

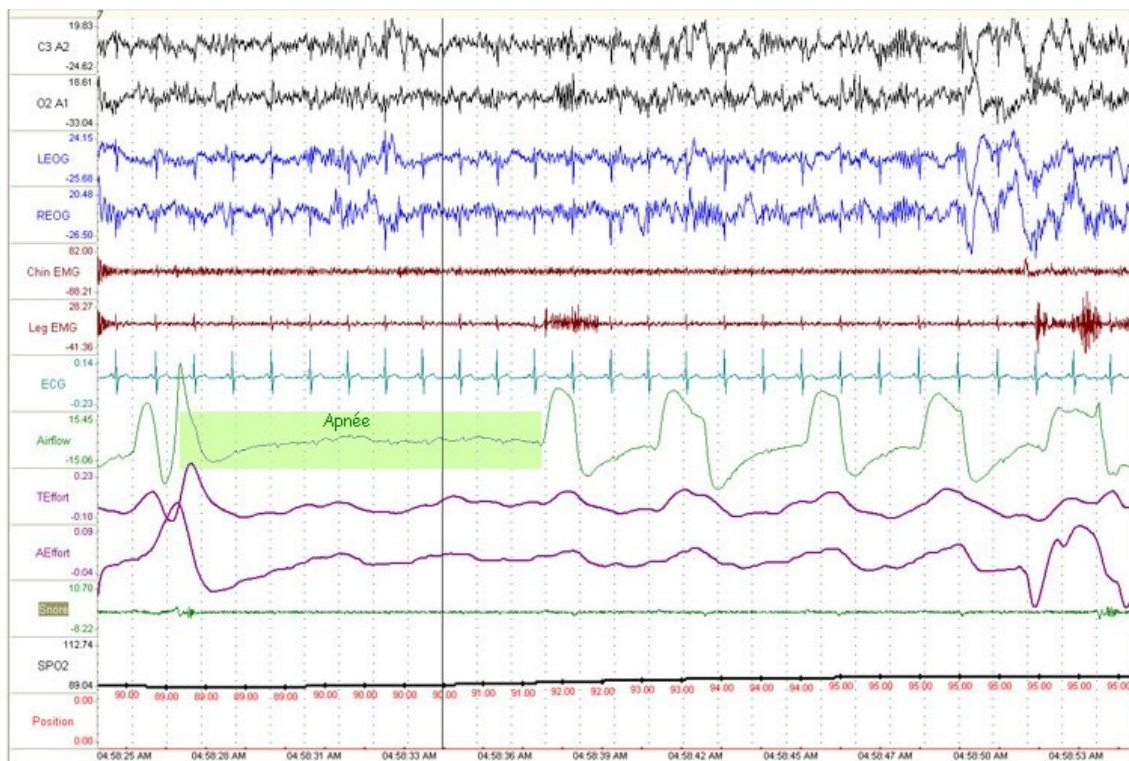


FIGURE 1.13 – Courbes polysomnographiques [réelles]



Les apnées sont ensuite classées en fonction des efforts respiratoires :

1. Une apnée est une apnée obstructive si elle est associée à une présence continue ou croissante des efforts respiratoires pendant toute la durée d'absence de flux respiratoire. (voir Figure 1.14)
2. Une apnée est une apnée centrale si elle est associée à une absence des efforts respiratoires pendant toute la période d'absence de flux respiratoire. (voir Figure 1.15)
3. Une apnée est une apnée mixte si elle est associée à une absence des efforts respiratoires dans sa première partie, suivie par une reprise des efforts respiratoires dans sa deuxième partie. (voir Figure 1.16)

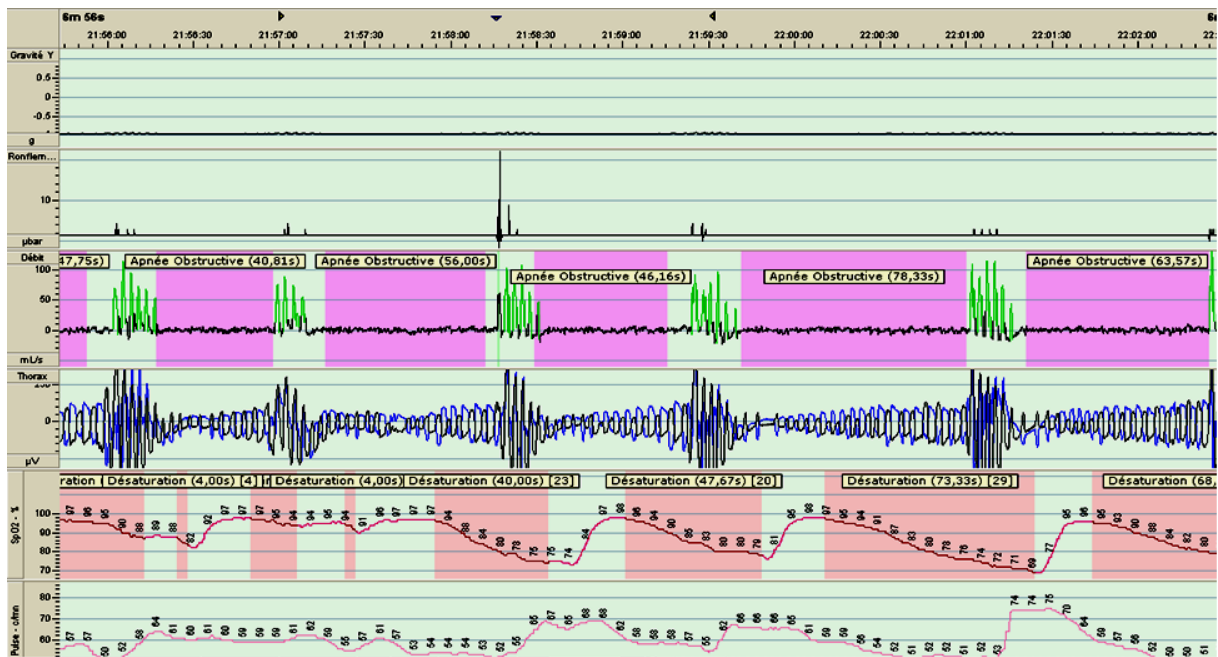


FIGURE 1.14 – Apnées obstructives - Courbes réelles

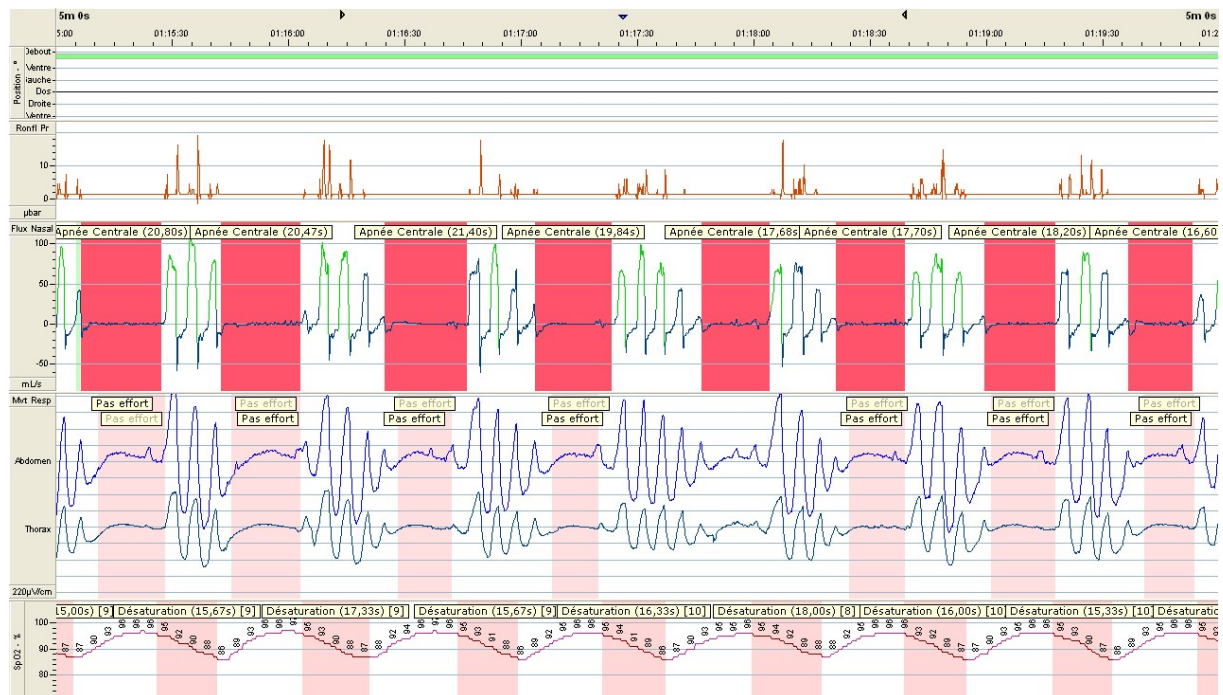


FIGURE 1.15 – Apnées centrales - Courbes réelles

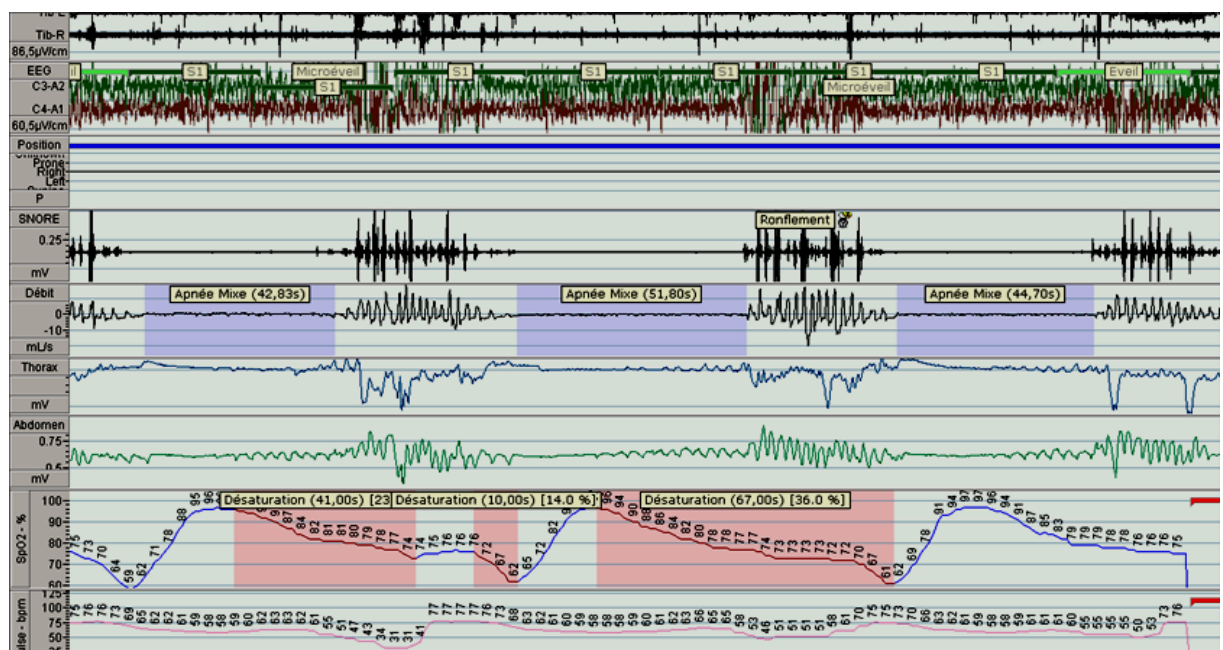


FIGURE 1.16 – Apnées mixtes - Courbes réelles

### Hypopnées

Il n'existe pas de consensus pour la définition des hypopnées. L'AASM propose deux séries de critères pouvant être utilisées comme définition d'une hypopnée [31] :

1. Première définition,
  - (a) Une hypopnée doit durer au moins 10 secondes.
  - (b) Il y a une baisse d'amplitude d'au moins 30% sur le signal de pression nasale pendant au moins 90% de la durée de l'hypopnée.
  - (c) Il y a une désaturation d'au moins 4%.
2. Deuxième définition,
  - (a) Une hypopnée doit durer au moins 10 secondes.
  - (b) Il y a une baisse d'amplitude d'au moins 50% sur le signal de pression nasale pendant au moins 90% de la durée de l'hypopnée.
  - (c) Il y a une désaturation d'au moins 3% et/ou que l'événement est associé à un micro-éveil.

La Figure 1.17 représente un schéma d'hypopnée enregistrée sur une courbe de flux respiratoire.

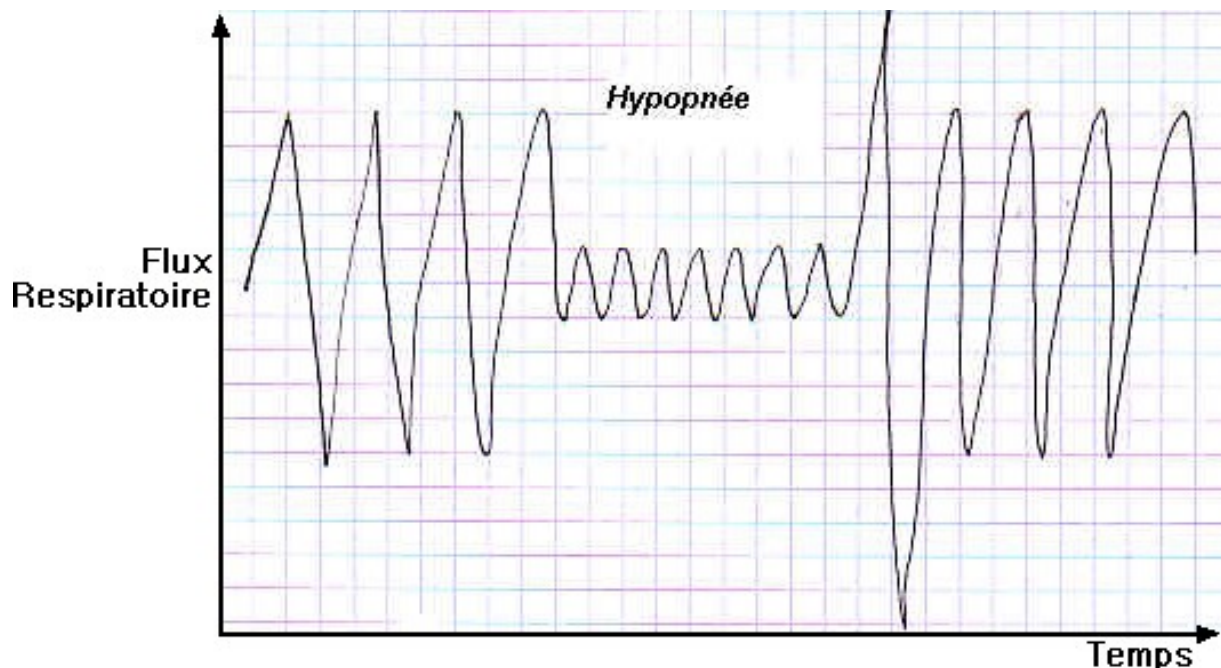


FIGURE 1.17 – Hypopnée - Schéma

La Figure 1.18 représente une hypopnée relevée sur des courbes réelles.

Les hypopnées peuvent également être classifiées en hypopnées obstructives ou centrales, mais cela nécessite d'avoir enregistré l'un des paramètres suivants :

- Pression œsophagienne qui relève la pression intrapleurale ;
- EMG diaphragmatique ou intercostal.

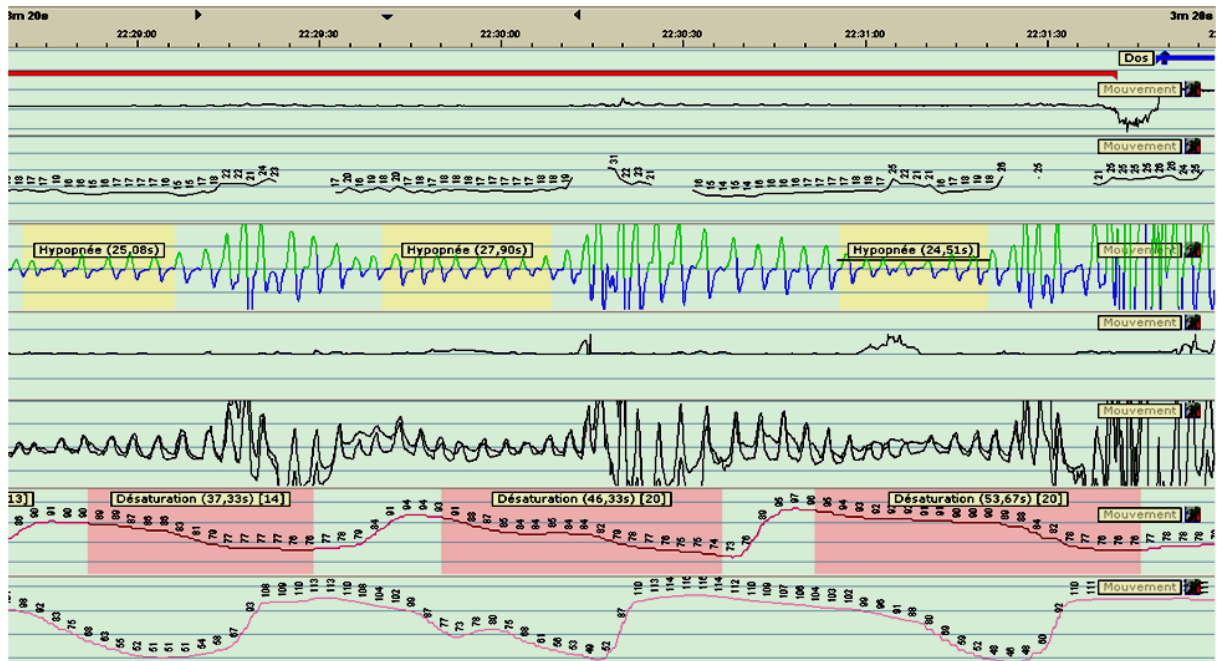


FIGURE 1.18 – Hypopnées - Courbes réelles

### Recherche de la nature spontanée ou liée des micro-éveils

Une fois que l'ensemble des événements « élémentaires » a été détecté, il faut mettre en évidence des liens de cause à effet entre les événements respiratoires et les micro-éveils. En effet, un micro-éveil peut survenir pour plusieurs raisons, parmi lesquelles la reprise de l'activité respiratoire. Il peut également être spontané. Il s'agit donc de décider, pour chaque micro-éveil, s'il est de nature spontanée ou s'il résulte d'un événement respiratoire.

Sur la figure 1.12, on peut observer une apnée, suivie d'un micro-éveil et d'une désaturation. Ce micro-éveil a pour cause l'apnée qui le précède et n'est donc pas spontané.

## 2.3 Diagnostic

Le diagnostic se fait, très grossièrement, par l'interprétation de deux indicateurs principaux :

- **Index d'Apnées-Hypopnées** : IAH ;

Il est défini comme la fréquence horaire moyenne des apnées et hypopnées, conformément à la formule suivante :

$$IAH = \frac{N(Apnees) + N(Hypopnees)}{TST}^{10}$$

L'IAH est l'indicateur synthétisant les événements symptomatiques du Syndrome d'Apnées du Sommeil survenus pendant l'enregistrement polysomnographique.

- le score sur l'échelle de somnolence d'Epworth (cf Figure 1.19) ;

Ce score évalue l'impact d'un mauvais sommeil sur la qualité de vie du patient, en évaluant sa propension à s'endormir dans différentes situations.

Une fois l'ensemble des courbes interprété, le médecin doit procéder au diagnostic. Pour cela, il s'aide de différents indicateurs. Le principal indicateur utilisé pour le diagnostic d'un Syndrome d'Apnées du Sommeil est l'Index d'Apnées-Hypopnées (IAH).

Bien sûr, il existe de nombreux autres indicateurs calculés depuis l'interprétation des courbes.



**Quelle a été la probabilité pour que vous vous assoupissiez ou que vous vous endormiez dans les conditions suivantes et non pas simplement parce que vous vous sentez fatigué?**

Pensez à votre façon de vivre habituelle. Même si vous ne vous êtes pas récemment trouvé(e) dans de telles situations, essayez d'imaginer comment celles-ci pourraient vous affecter.

Pour répondre, utilisez l'échelle suivante en choisissant le chiffre le plus approprié pour chaque situation:

0: pas de risque de s'assoupir

1: petite chance de s'assoupir

2: possibilité moyenne de s'assoupir

3: grande chance de s'assoupir

Situation	Chance de s'assoupir			
Assis en lisant	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
En regardant la télévision	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
Assis inactif dans un lieu public (par exemple au théâtre ou lors d'une réunion)	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
Comme passager d'une voiture roulant depuis 1 heure sans arrêt	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
En s'allongeant l'après-midi pour une sieste, lorsque les circonstances le permettent	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
Assis en discutant avec un proche	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
Assis tranquillement après un repas sans alcool	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
Au volant d'une voiture immobilisée depuis quelques minutes dans un embouteillage	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3

**Total des points:**

**Evaluation**

0–10 points:

11–14 points:

15 points et plus:

absence de somnolence

somnolence légère

somnolence sévère

FIGURE 1.19 – Échelle de somnolence d'Epworth

Dans [31], l'AASM recommande d'inclure les paramètres suivants dans chaque compte-rendu polysomnographique :

- contexte médical du patient,
  - âge ;
  - taille ;
  - poids ;
  - antécédents médicaux ;
  - motif de consultation ;
  - score sur l'échelle de somnolence d'Epworth.
- techniques d'enregistrements,
  - liste des voies enregistrées ;
  - heure de début d'analyse ;
  - heure de fin d'analyse.
- structure du sommeil,
  - période de Sommeil Total (PST)<sup>11</sup> ;
  - temps de Sommeil Total (TST)<sup>12</sup> ;
  - index d'efficacité du sommeil : TST/PST.
- architecture du sommeil,
  - organisation des cycles du sommeil et des temps d'éveil au cours de la nuit ;
  - organisation des stades de sommeil dans les différents cycles ;
  - hypnogramme<sup>13</sup> ;
  - répartition du temps de sommeil par stade de sommeil ;
  - répartition du temps de sommeil par position.
- latence et continuité,
  - latence d'endormissement<sup>14</sup> ;
  - latence du sommeil paradoxal<sup>15</sup> ;
  - nombre d'éveils de plus de 60 secondes.
- index de micro-éveils,
- index des événements respiratoires,
  - index global des événements respiratoires ;
  - index d'Apnées-Hypopnées ;
  - type d'événement majoritaire ;
  - durée moyenne des événements ;
  - durée de l'événement le plus long ;
  - index des événements respiratoires par stade de sommeil ;
  - index des événements respiratoires par position.

---

11. Durée entre la première minute de sommeil et le réveil définitif

12. Durée pendant laquelle le patient est endormi

13. Représentation graphique temporelle des stades de sommeil successifs pendant la nuit

14. Durée écoulée entre l'extinction des lumières et l'endormissement

15. Durée écoulée entre l'endormissement et l'apparition du premier stade de sommeil paradoxal

- données de saturation,
  - saturation moyenne à l'éveil ;
  - saturation moyenne lors de l'enregistrement ;
  - saturation minimale ;
  - pourcentage du temps passé avec une saturation inférieure à 90% ;
  - index de désaturations.
- données cardiaques ;
  - fréquence cardiaque moyenne ;
  - intervalle de répartition des fréquences cardiaques.

## 2.4 Traitement

Le traitement habituel d'un Syndrome d'Apnées du Sommeil important (en France,  $IAH > 30$  ou index de micro-éveils  $> 10$ ) est la Pression Positive Continue (PPC) délivrée par une machine, au travers d'un masque que le patient doit porter quand il dort (cf Figure 1.20)

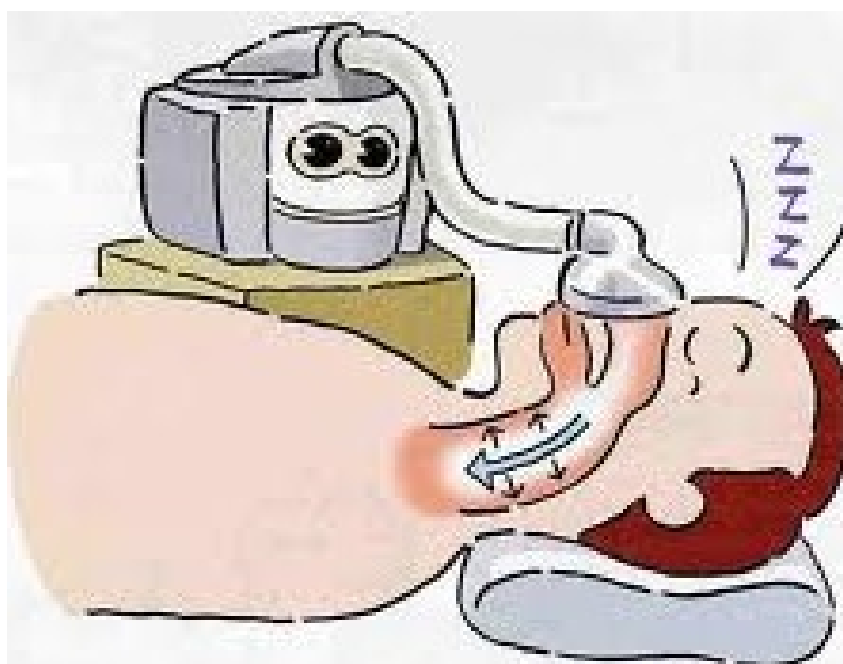


FIGURE 1.20 – Traitement du SAS par Pression Positive Continue

## 3 Variabilité inter- et intra-experts des interprétations visuelles

Le Rapport sur le thème du sommeil [26] remis en décembre 2006 par le Dr Jean-Pierre GIORDANELLA, médecin en santé publique, Directeur de la prévention à la Caisse Primaire d'Assurance Maladie (CPAM) de Paris, à Xavier BERTRAND, ministre de la Santé et des Solidarités, recommande de ne pas se fier aux analyses automatiques souvent proposées dans les matériels actuels, et, lorsqu'on y a recours, qu'elles soient « validée[s] et



éventuellement modifiée[s] par un médecin dont la signature l'engage dans cet acte médical. » Ceci montre bien qu'il est pour le moment considéré que les analyses automatiques de données polysomnographiques ne sont pas fiables. Différentes études ont abouti aux mêmes conclusions :

- Dans [22], Raffaele FERRI, médecin dans le *Department of Neurology I.C* du *Sleep Research Center* à l'*Oasi Institute (IRCCS)* à Troina en Italie a Comparé les résultats inter-scoreurs d'une analyse visuelle de données polysomnographiques. Elle a montré qu'il existait une grande variabilité dans les résultats entre 9 groupes de *scoreurs*. Chaque groupe est constitué de 1 à 2 médecins du même service ; ils se répartissent l'interprétation des courbes. Les résultats de la première colonne du tableau 1.2 donnent l'échelle des concordances obtenues en comparant les interprétations deux à deux.
- Masaaki SUZUKI, du *Department of Otolaryngology* à la faculté de médecine de l'Université de Teikyo à Tokyo a lui aussi étudié dans [72] la concordance entre interprétations visuelles de signaux polysomnographiques. Il a obtenu les résultats exposés dans la deuxième colonne du tableau 1.2. Les résultats donnent l'échelle des concordances obtenues en comparant les interprétations deux à deux.
- Dans [57], Robert G NORMAN du *Department of Pulmonary and Critical Care Medicine* à la faculté de médecine de New-York a également étudié cette concordance entre 5 médecins sur 62 polysomnographies . Les résultats en troisième colonne du tableau 1.2 donnent le pourcentage d'époques, pour chaque stade de sommeil, où l'ensemble des scoreurs sont d'accord parmi toutes les époques où au moins l'un des scoreurs a associé ledit stade de sommeil.
- Heidi DANKER-HOPFE, du *Department of Psychiatry and Psychotherapy* à l'*Hôpital de la Charité* à Berlin a étudié en 2004, puis en 2009, la concordance d'interprétation visuelle des stades de sommeil entre experts humains issus de laboratoires différents. En 2004, des experts de 8 laboratoires européens différents ont analysé les polysomnographies de 98 patients différents ; chaque polysomnographie étant analysée par 2 laboratoires. La concordance a été de 76,8% ( $\kappa = 0,6816$ ). En 2009, des experts de 3 des laboratoires ayant participé à l'étude précédente ont réinterprété 72 polysomnographies (issues de l'ensemble utilisé pour l'étude de 2004), en suivant cette fois ci les nouvelles règles de l'AASM publiées dans [31] en 2006. Sur ce sous-ensemble, la concordance vaut 80,6% ( $\kappa = 0,68$ ) en utilisant les règles de Rechtschaffen & Kales [67], et 82,0% ( $\kappa = 0,76$ ) en utilisant les règles de l'AASM [31].
- L'étude principale concernant les divergences d'interprétation est celle réalisée par Michael H. Silber dans [69]. Elle a été utilisée comme référence par l'AASM pour définir les nouvelles règles. Différentes publications ont été examinées pour analyser la concordance inter- et intra-scoreurs. Les résultats figurent en gras dans les quatrième et cinquième colonnes du tableau 1.2

Certains travaux portent sur la comparaison de l'interprétation des événements respiratoires. Dans [12], Pierre-Olivier BRIDEVEAUX du service de Pneumologie à l'Hôpital de Lausanne a également étudié les concordances d'interprétation entre, d'une part, 8 analyses visuelles faites par des experts « humains » sur 11 polygraphies ventilatoires, et, d'autre part, avec une analyse automatique. Au final, les IAH des différentes analyses, humaine et automatique, sont comparées. Des rapports de 1 à 4 sont obtenus entre les IAH issus d'interprétations visuelles. L'analyse automatique a, quant à elle, tendance à sous-estimer la quantité d'hypopnées, et, par conséquence, à sous-évaluer l'IAH.

TABLEAU 1.2 – Concordance d'interprétation visuelles de stades de sommeil entre scoreurs « humains »

Stade de sommeil	Ferri <i>et al</i>	Suzuki <i>et al</i>	Norman <i>et al</i>	Silber <i>et al</i> (inter-scoreur)	Silber <i>et al</i> (intra-scoreur)
W	47,5-100%	17,2-92,7%	47,8%	<b>68-89%</b>	<b>89-93%</b>
R	62,9-100%	84-97,9%	42,3%	<b>78-94%</b>	<b>72-88%</b>
N1	33,3-96,3%	34,6-81,9%	3,6%	<b>23-74%</b>	<b>18-42%</b>
N2	39,9-99,1%	16,2-89,2%	35,4%	<b>79-90%</b>	<b>76-85%</b>
N3	32,3-100%	0-76,5%	18,6%	<b>69%</b>	<b>55-75%</b>

D'autres travaux ont étudié la détection des micro-éveils. Michael J. DRINNAN des *Departments of Regional Medical Physics and Respiratory Medicine* au *Freeman Hospital* à Newcastle Upon Tyne a demandé, dans [20], à des experts issus de 14 laboratoires différents, d'analyser visuellement 90 événements EEG issus des polysomnographies de 10 patients, afin de décider s'il s'agissait, ou non, de micro-éveils. Au final, 23 à 53 des 90 événements ont été étiquetés comme micro-éveils, avec une concordance moyenne globale  $\kappa = 0,47$ .

Enfin, Nancy A COLLOP, de la *Division of Pulmonary/critical Care Medicine* du centre médical de l'Université du Mississippi à Jackson aux États-Unis, a comparé dans [16] les IAH après des interprétations complètes faites par des experts humains. 11 médecins ont fait l'interprétation complète de 11 polysomnographies différentes. La corrélation entre les IAH des différents experts vaut seulement  $\kappa = 0,24$ , ce qui indique une faible concordance [35].

Toutes ces études montrent bien que le problème de l'interprétation visuelle des données polysomnographiques a une variabilité inter- et intra-experts non négligeable. C'est une information importante pour l'évaluation de nos résultats (et donc de notre méthode) car ceci signifie qu'il n'existe pas de référence absolue pour les différentes interprétations des signaux polysomnographiques. Néanmoins, l'interprétation visuelle par un expert est aujourd'hui la plus fiable et doit être utilisée comme référence.

## 4 Pertinence d'un système de support à la décision

Peter ANDERER du *Departments of a Psychiatry and Psychotherapy* du *Center for Brain Research* de la faculté de médecine de Vienne, membre du *Siesta Group Schlafanalyse GmbH*, a montré dans [5] que la concordance d'interprétation des signaux polysomnographiques pouvait être améliorée avec l'utilisation d'une analyse automatique. En effet, 72 enregistrements polysomnographiques ont été interprétés d'une part, visuellement par 2 experts humains, et, d'autre part, par une analyse automatique. Cette analyse automatique a ensuite été corrigée par 2 experts humains. Les résultats sont présentées dans le tableau 1.3.

TABEAU 1.3 – Concordance d'interprétations de signaux polysomnographiques entre analyses visuelles, automatiques et semi-automatiques

	Interprétations visuelles	Analyses automatiques	Analyses semi-automatiques
Accord	82%	81%-82%	99%
$\kappa$	0,76	0,75-0,76	0,99

## 5 Conclusion

Le Syndrome d'Apnées du Sommeil (SAS) est un cadre de travail qui convient tout à fait aux méthodes de Fusion Symbolique que nous désirons tester. En effet, cette pathologie est un véritable problème de santé publique car sa prévalence est estimée à 2% des femmes d'âge moyen et à 4% des hommes d'âge moyen [80]. De plus, l'examen de référence aujourd'hui est la polysomnographie, qui consiste en l'enregistrement simultané de différents paramètres physiologiques au cours du sommeil, donc souvent sur plus de 10 heures. Chacune de ces sources apporte une information qui va contribuer au diagnostic final. Leur interprétation est une tâche chronophage et fastidieuse entraînant des divergences d'un expert à l'autre. Le Syndrome d'Apnées du Sommeil a diverses conséquences : tout d'abord une somnolence diurne excessive, démontrée comme étant l'une des causes majeures de survenue d'accidents de la route ou du travail, mais aussi un surcroît de morbidité et de mortalité cardio et cérébro-vasculaires. De plus, selon le rapport du Dr GIORDANELLA [26], le SAS serait sous diagnostiqué, et seuls 10 à 15 % des patients seraient pris en charge en France. Ce même rapport reconnaît que ceci s'explique en partie parce que le diagnostic du SAS est chronophage et recommande d'utiliser un outil de diagnostic simple. C'est donc un contexte particulièrement motivant par les enjeux médicaux et sociétaux qui lui sont associés.

Le support au diagnostic du Syndrome d'Apnées du Sommeil est également un défi important pour la fusion symbolique. En effet, après avoir collecté, visuellement ou automatiquement, l'ensemble des informations nécessaires à l'établissement d'un diagnostic, il s'agit de les fusionner afin de mettre en évidence les liens physiologiques qui existent entre deux événements physiopathologiques. Souvent, les événements eux-mêmes sont issus de la fusion d'observations de caractéristiques survenant simultanément, ou presque, sur différentes courbes. Les concepts à utiliser et de nombreuses règles de fusion sont déjà définies par l'AASM dans [31]. Il s'agit donc d'un domaine qui se prête favorablement à la mise en place d'un système de support à la décision reposant sur une fusion symbolique.



# Chapitre 2

## Reconnaissance des stades de sommeil

Les travaux de cette thèse sont centrés sur l'utilisation de la fusion symbolique pour l'interprétation automatique des données polysomnographiques. La fusion sert à mettre en relation des « objets » afin qu'ils soient sémantiquement plus riches et à augmenter le niveau de connaissance.

La fusion a longtemps été appliquée aux courbes numériques. Nous pensons que, appliquée aux données symboliques, la fusion peut obtenir de meilleurs résultats, et, dans le cas des données polysomnographiques, d'atteindre un niveau de connaissances permettant un support au diagnostic efficace pour le Syndrome d'Apnées du Sommeil.

La fusion symbolique consiste en la fusion d'informations présentées sous un formalisme symbolique suivant des règles associées à une stratégie prédéfinie. Pour cela, il faut disposer, au préalable, d'éléments d'information unitaires représentés selon le formalisme choisi.

Dans notre cas, il faut procéder à une étape préliminaire de recueil de ces informations. Elles sont à extraire de l'enregistrement polysomnographique, c'est-à-dire que, pour chaque patient, il faut extraire des courbes polysomnographiques (au format numérique) un ensemble d'informations interprétées sémantiquement et les représenter dans le même formalisme.

Ces informations se divisent en deux groupes : d'une part, les événements constitutifs du Syndrome d'Apnées du Sommeil tels que les apnées ou les désaturations, et, d'autre part, les conditions environnementales ou contextuelles de survenue des événements précédemment cités, notamment les stades de sommeil. C'est à leur reconnaissance que nous souhaitons nous confronter dans un premier temps.

Nous avons choisi ici d'utiliser comme formalisme les graphes conceptuels pour représenter l'information. Les règles de fusion s'inspirent largement des recommandations médicales officielles à utiliser pour le diagnostic du Syndrome d'Apnées du Sommeil.

### 1 Données sources

Nous avons eu la chance, pour ce doctorat, de pouvoir travailler sur 215 enregistrements polysomnographiques, tous intégralement interprétés et diagnostiqués par un méde-

cin spécialiste du sommeil. Au total, ceci représente près de 3200 heures d'enregistrement.

La plupart des enregistrements polysomnographiques à notre disposition était constituée de :

- 3 EEG correspondant aux dérivations suivantes (voir figure 2.1) :
  - C4-A1 ;
  - C3-A2 ;
  - O1-A2 ;

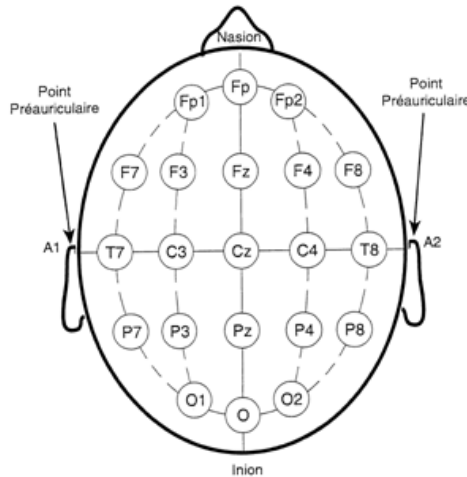


FIGURE 2.1 – Positions sur le scalp pour la pose des capteur EEG

- 2 EOG (voir figure 2.2) :
  - EOG-L (œil gauche) ;
  - EOG-R (œil droit) ;

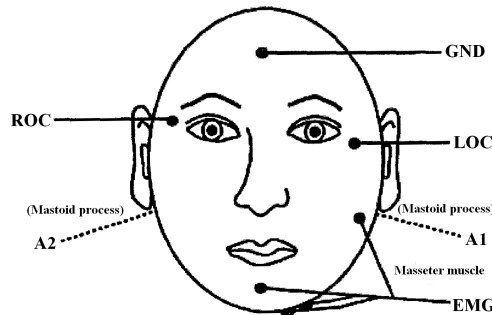


FIGURE 2.2 – Positions des capteurs faciaux

- 1 EMG submentonnier ;
- 1 canule nasale ;
- 2 ceintures thoraco-abdominales ;
- 1 oxymètre mesurant la saturation en Oxygène ( $SpO_2$ ) et le pouls ;
- 1 ECG ;
- 2 EMG tibiaux ;
- 1 capteur de position du corps.

Certains ont en plus une thermistance buccale pour mesurer d'éventuelles respirations buccales.

Ces enregistrements ont été effectués avec le système ®EMBLA et interprétés par un expert médical avec le logiciel <sup>TM</sup>Somnologica de la société ResMed qui nous a envoyé le fichier de description de format [51] afin de pouvoir décoder les fichiers et importer les données.

Sur les figures 2.3 et 2.4 sont représentés des aperçus d'écran du logiciel <sup>TM</sup>Somnologica afin de voir l'environnement de travail de l'expert médical pour l'analyse visuelle, mais aussi comment sont annotés les événements repérés visuellement par le médecin (voir Figure 2.4).

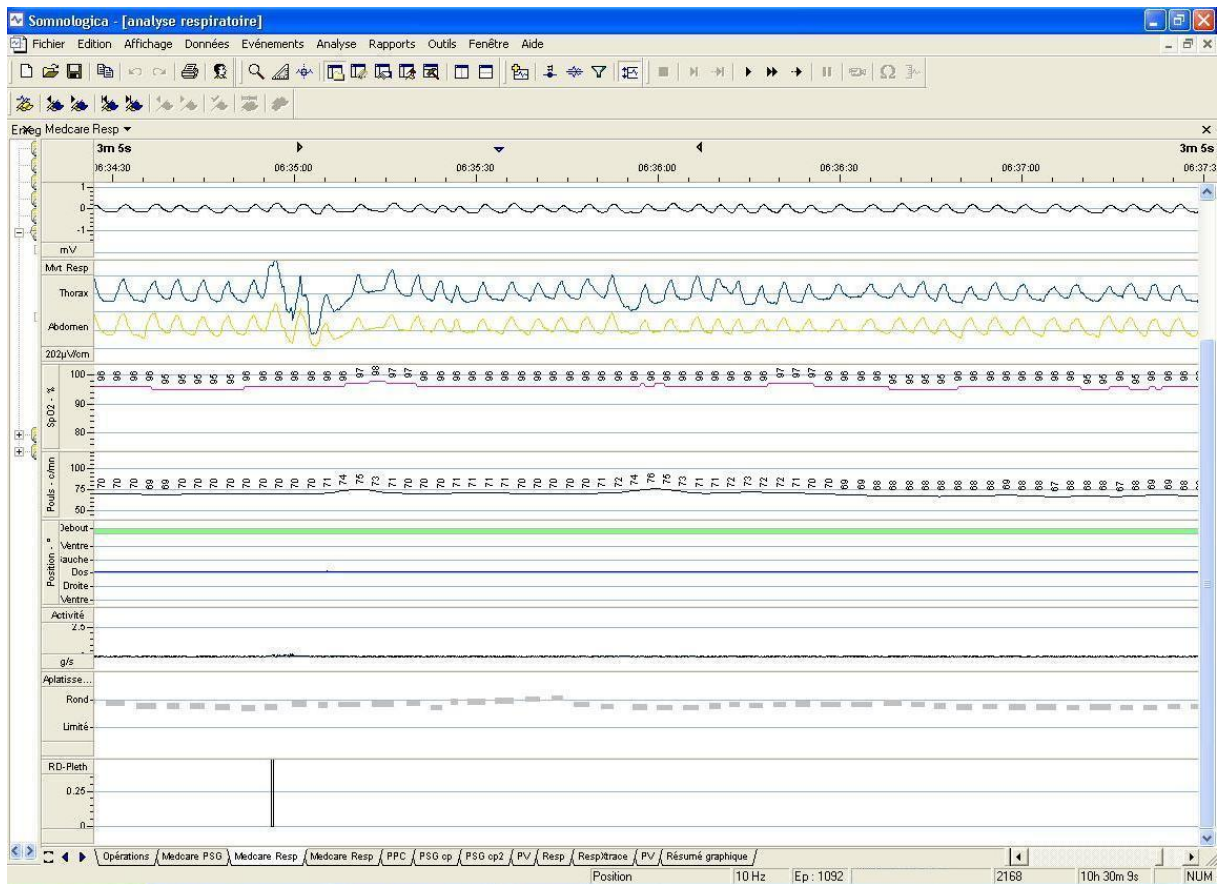
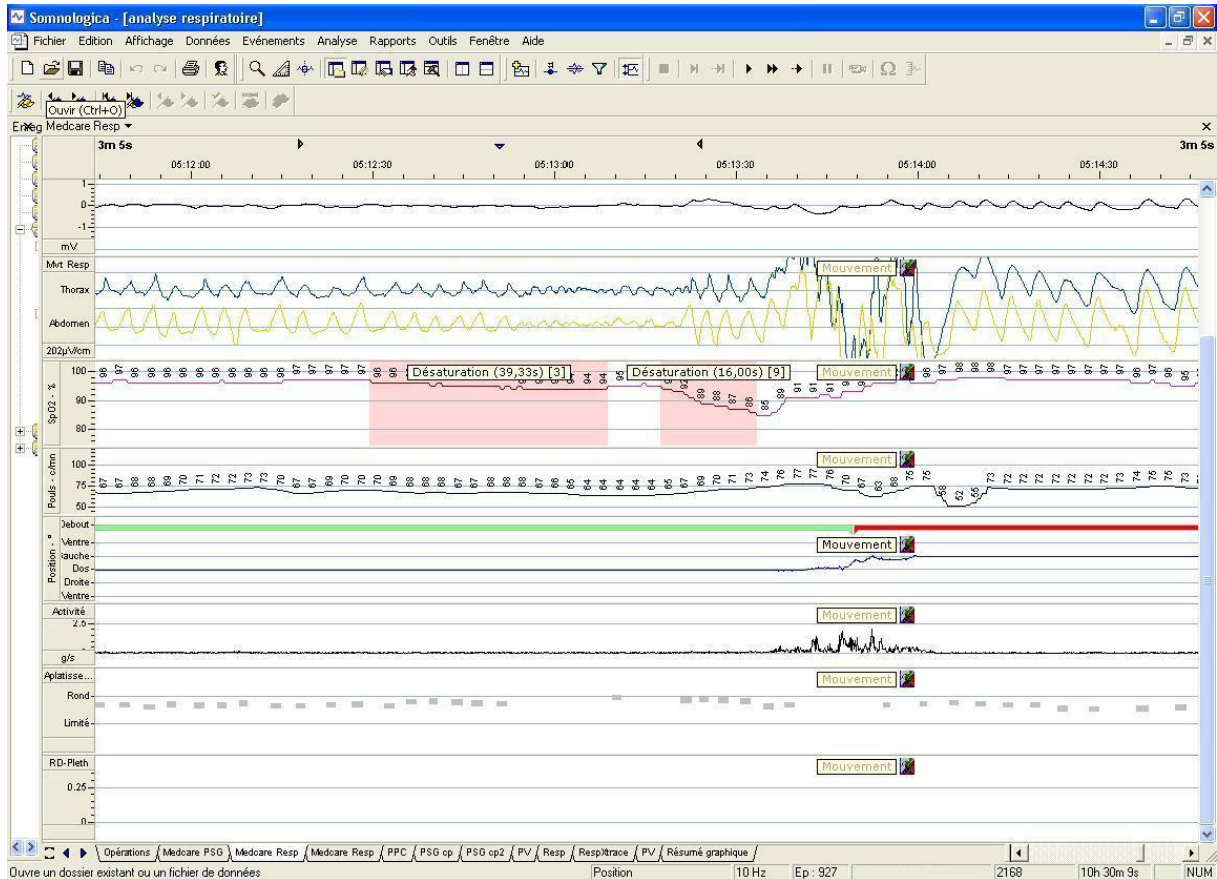


FIGURE 2.3 – Logiciel <sup>TM</sup>Somnologica

À l'hôpital Tenon, chaque patient qui passe une polysomnographie remplit à l'issue de la nuit un questionnaire où il indique, entre autres, l'heure de coucher et l'heure de lever. Ces informations sont utilisées ensuite par le médecin qui va restreindre son analyse précise à cette période, ne regardant que rapidement le reste de l'enregistrement, sans l'annoter. Ce questionnaire ne nous a pas été donné, mais il est possible de retrouver la période d'analyse qui correspond exactement à la période de scoring.

Les données sources sont donc constituées de 215 dossiers, un pour chaque patient. Chaque dossier patient contient un fichier embla (.ebm) par voie enregistrée et un fichier au format XML qui est le rapport des événements scorés par l'expert médical.



FIGURE 2.4 – Logiciel <sup>TM</sup>Somnologica avec 2 désaturations relevées par l'expert

## 2 Détection automatique des stades de sommeil par arbre binaire de décision

La fusion symbolique nécessite de disposer d'un ensemble d'informations représentées sous un format symbolique. Il est donc nécessaire, dans un premier temps, de procéder à l'interprétation des stades de sommeil pour chacune des époques des signaux neurophysiologiques enregistrés lors de la polysomnographie. Dans cette première méthode, nous allons tester un classifieur qui fusionne un mélange de données numériques et symboliques par un arbre binaire de décision.

Ce travail a fait l'objet d'une publication [75], présentée oralement à la *13th conference on Artificial Intelligence in Medicine (AIME'11)* qui s'est tenue à Bled en Slovénie du 2 au 6 Juin 2011.

### 2.1 État de l'art

Le problème de la stadification automatique du sommeil est un problème auquel se sont confrontées de nombreuses équipes de recherche. L'équipe de Christian BERTHOMIER et Jacques PRADO de l'Hôpital Henri MONDOR à Créteil (AP-HP) ont proposé en 1999 une analyse par bancs de filtres sur une voie EEG [11].



En 2001, Jung-Su OH du *Institute of Biomedical Engineering* de la faculté de médecine de Séoul, a présenté dans [58] une méthode de reconnaissance des stades de sommeil utilisant un système à modules interconnectés :

- un module de traitement du signal ;
- un module de raisonnement par règles ;
- un module de raisonnement par cas.

En 2002, Masaaki HANAOKA, de la *Graduate School of Engineering* à l'Université *Yamanashi* à Kofu au Japon a proposé une méthode utilisant un arbre de décision utilisant des paramètres issus de la reconnaissance d'ondes par extraction de paramètres non fréquentiels ( sur 4 voies EEG, 2 voies EOG, 1 voie EMG et 1 voie ECG chez des sujets sains [30].

En 2003, Thomas PENZEL, de l'Hôpital de l'Université Philipps, à Marburg en Allemagne, a publié dans [62], des travaux de reconnaissance des stades de sommeil chez des sujets atteints d'un syndrome d'Apnées du Sommeil, modéré ou sévère, par analyse des fluctuations détendancées (*Detrended Fluctuation Analysis*) d'une part et par analyse spectrale de la série temporelle des intervalles RR d'autre part, obtenue à partir d'enregistrements ECG.

En 2005, Nizar KERKENI, de l'équipe *Cortex* au LORIA à Vandœuvre-lès-Nancy en France, a publié une méthode utilisant un réseau de neurones sur des paramètres issus d'une analyse spectrale d'un signal EEG. Les tests ont été faits sur un seul enregistrement polysomnographique [32].

En 2007, Tarik Al-ani, de l'École Supérieure d'Ingénieurs en Électronique et Électrotechnique de l'Université Paris-Est à Noisy-le-Grand a proposé une méthode de stadification automatique du sommeil par analyse du signal électrocardiographique en 2 étapes : extraction de paramètres en utilisant des méthodes d'analyse des fluctuations détendancées et d'analyse de la variabilité du rythme cardiaque. Ces paramètres sont ensuite injectés dans un réseau de neurones. Ces travaux sont publiés dans [4].

En 2007, Christian BERTHOMIER, de *PHYSIP SA*, et Xavier DROUOT, du Groupe Hospitalier Henri Mondor à Créteil (AP-HP), ont présenté dans [10] des travaux de reconnaissance des stades de sommeil sur une voie EEG, en utilisant un système appelé ASEEGA dont l'algorithme est découpé en 3 parties, puis en sous-parties :

- preprocessing,
  1. rejet des artefacts ;
  2. réglage des bandes de fréquences ;
  3. bandes de filtres non uniformes.
- analyse,
  1. extraction de paramètres spectraux et temporels ;
  2. détection de microstructures ;
  3. identification grossière des réveils ;
  4. identification grossière des mouvements oculaires rapides.

- classification,
  1. classification floue ;
  2. lissage par règles contextuelles ;
  3. fusion des stades de sommeil.

En 2009, Luay A. FRAIWAN, du département d'ingénierie biomédicale de l'Université de Sciences et de Technologie à Irbid en Jordanie, a publié dans [23] des travaux d'analyse automatique des stades de sommeil extrayant des paramètres suite à une décomposition en ondelettes du signal EEG, puis en injectant ces paramètres dans un arbre de décision. Les travaux ont été effectués sur 32 enregistrements polysomnographiques et ont abouti à une concordance de 75% avec l'expert.

En 2010, Salih Günes, du *Electrical and Electronics Engineering Department*, de l'Université de Selçuk en Turquie a présenté dans [28] une méthode d'analyse automatique des stades de sommeil en 3 temps :

1. Dans un premier temps, 129 paramètres sont extraits par une analyse spectrale de Welch sur un signal EEG. L'analyse spectrale de Welch est une méthode basée sur la transformée de Fourier rapide (Fast Fourier Transform ou *FFT*). L'ensemble des 129 paramètres est réduit à 4 paramètres en calculant leur minimum, leur maximum, leur moyenne et leur écart-type.
2. Chaque paramètre est ensuite pondéré par une méthode de pondération de paramètres basée sur l'algorithme des *k-means* (*k-Means clustering based feature weighting* (*KMCFW*)).
3. La classification finale utilise l'algorithme des K plus proches voisins.

Cette méthode permet d'obtenir une concordance globale de 82,15%. Sur les 4196 époques disponibles, 3400 ont été utilisées pour l'ensemble de test.

En 2011, Haoyu MA, de la *School of information science and engineering* à l'Université de Lanzhou en Chine ont proposé dans [46] une méthode de reconnaissance automatique des stades de sommeil par association d'un réseau de neurones et d'un arbre de décision.

## 2.2 Stadification automatique par arbre binaire de décision

Un arbre de décision est un outil d'aide à la décision qui accepte en entrée une situation ou un objet décrit par un ensemble d'attributs et retourne une « décision », la valeur de sortie prédite. Les attributs d'entrée, de même que la valeur de sortie, peuvent être discrets ou continus. Un arbre de décision atteint une décision en exécutant un ensemble de tests successifs. Chaque nœud de l'arbre correspond à un test de l'une des propriétés sur un ou plusieurs attributs ; les branches sortant du nœud correspondent aux différentes valeurs, ou ensembles de valeurs, du test. Dans le cas d'un arbre binaire, chaque test est un booléen qui teste si la propriété est vraie ou fausse. Les feuilles de l'arbre indiquent la valeur à retourner. L'apprentissage d'une fonction continue s'obtient par régression ; l'apprentissage d'une fonction discrète s'obtient par classification.

Dans notre cas, nous allons utiliser un arbre binaire de décision à valeurs d'attributs discrètes.

## 2.3 Hypothèses de travail

Pour effectuer cette analyse automatique des stades de sommeil, nous avons travaillé avec 3 hypothèses principales :

1. Les époques vont être analysées indépendamment les unes des autres ; Ceci signifie que l'aspect séquentiel des stades de sommeil, bien connu des experts et inévitablement utilisé par tout « scoreur » humain, va être ici ignoré. Nous avons en fait testé des algorithmes d'apprentissage d'arbre binaire de décision sous WEKA en utilisant, comme l'un des attributs de description des données en entrée, le stade de sommeil de l'époque précédant celle analysée. Même si les résultats des tests de ces algorithmes sur l'ensemble de test sous WEKA atteignent des taux d'accord avec l'expert supérieurs à 90%, cet attribut prend un poids relatif aux autres attributs beaucoup trop important, ce qui entraîne une forte dépendance. En effet, quel que soit le stade de sommeil affecté à une époque  $N$ , alors le plus probable pour l'époque  $N+1$  est de rester dans le même stade de sommeil. En pratique, un tel classifieur donne des résultats très mauvais : en effet, au début de l'enregistrement, le patient est éveillé, et selon les résultats de ce classifieur, reste, en conséquence, éveillé tout au long de l'enregistrement.

2. Nous allons décrire les courbes d'entrée par des attributs essentiellement fréquentiels. En effet, les règles d'interprétation des stades de sommeil définies par l'AASM reposent essentiellement sur des paramètres fréquentiels (bandes de fréquences sur les enregistrement EEG caractéristiques de chaque stade de sommeil) ou interprétables dans l'espace fréquentiel (mouvements oculaires des yeux). Par cette hypothèse, nous excluons l'utilisation de paramètres indicatifs de la présence ou l'absence des patterns courts significatifs d'un stade de sommeil (complexes-K, fuseaux du sommeil ou pointes vertex) qui nécessiteraient une étape supplémentaire préalable de reconnaissances de ces éléments, ce qui est une étape difficile pour laquelle il n'existe pas à ce jour d'algorithme suffisamment fiable.

Afin de valider notre hypothèse, nous avons confronté les spectrogrammes de l'enregistrement EEG de 4 patients à leurs hypnogrammes. Un spectrogramme est un diagramme associant à chaque instant  $t$  son spectre de fréquence. Les résultats sont donnés dans la figure 2.5.

Il apparaît clairement une corrélation visuelle entre les couleurs du spectrogramme et l'hypnogramme, représenté en noir, ce qui conforte l'hypothèse énoncée. Ceci a également été confirmé par les travaux de Ahmad Taufiq MOHAMAD [55].

3. La dernière hypothèse de travail est l'utilisation de paramètres locaux en complément des paramètres globaux décrivant l'époque dans son intégralité. En effet, nous pensons que pour déterminer le stade de sommeil d'une époque donnée, il faut regarder l'allure globale des courbes, mais aussi des phénomènes locaux, caractéristiques d'un stade de sommeil. Les théories de Fourier montrent qu'un signal peut être décomposé comme une somme de sinusoïdes. Les signaux polysomnographiques sont complexes et la transformée de Fourier n'est pas simple à analyser ;

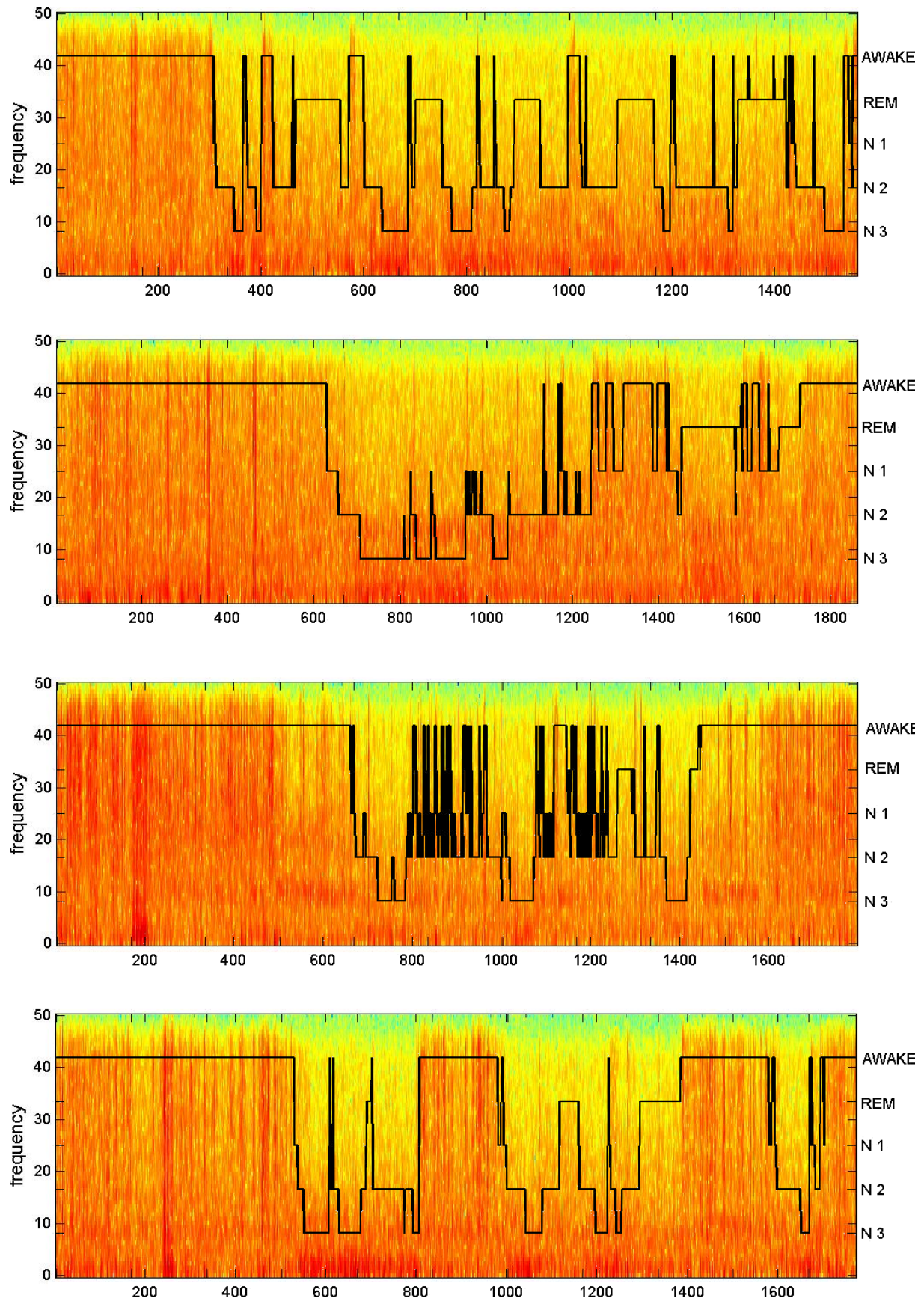


FIGURE 2.5 – Comparaison Spectrogramme / Hypnogramme



nous pensons que chaque époque renferme des fenêtres locales d'une seconde où le signal est plus proche des caractéristiques recherchées, ce qui peut aider à la décision.

Sur la figure 2.6 sont représentés les graphiques des transformées de Fourier de signaux EEG pour 6 époques avec des stades de sommeil différents ; les courbes d'éveil étant très différentes selon l'activité cérébrale, deux graphiques ont été tracés pour illustrer le stade W (Wakefulness) avec une époque en début d'enregistrement, avant extinction de la lumière, et une autre après que le patient se soit couché. En jaune sont représentées les bandes de fréquences du signal EEG caractéristiques de chaque stade de sommeil selon les règles de l'AASM [31]. Comme nous l'avons vu dans la partie 2.2 du chapitre 1, l'AASM ne définit plus de bande de fréquences spécifiques du stade N2 [31], ce qui explique l'absence de bande jaune sur le graphique correspondant.

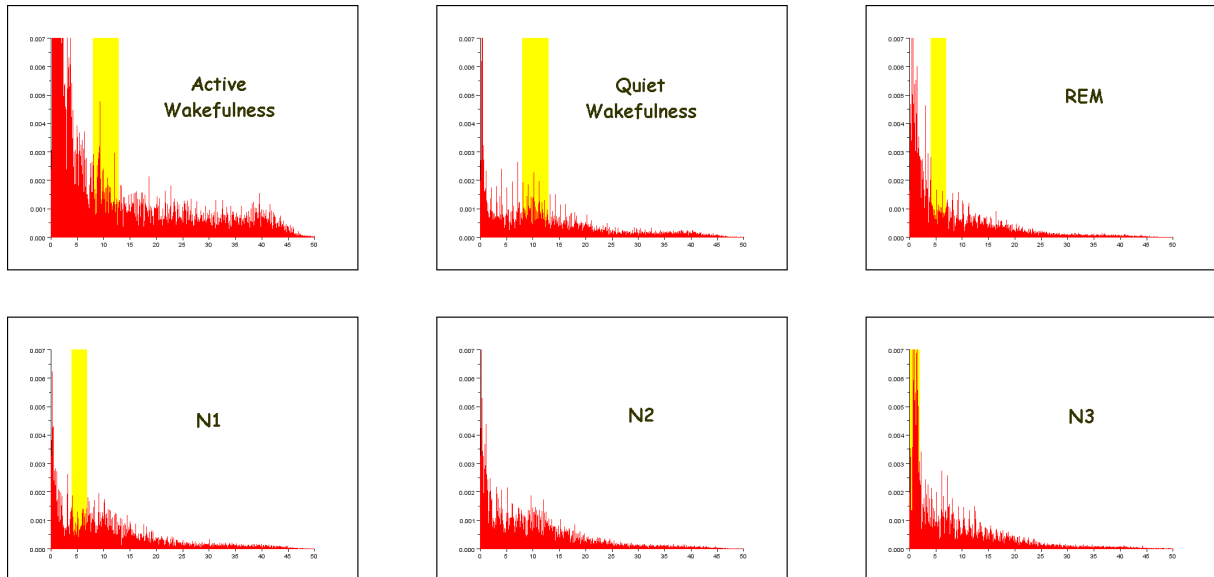


FIGURE 2.6 – FFT par stade de sommeil

Sur la figure 2.7 est représentée l'énergie du signal par stade de sommeil et par bancs de fréquences. Sur le premier graphique, on peut voir l'hypnogramme du patient selon l'expert, c'est-à-dire la succession des stades de sommeil tout au long de la nuit. Sur le deuxième graphique, on peut voir l'évolution de l'énergie du signal brut, c'est-à-dire non filtré. Sur le troisième graphique est représentée l'évolution de l'énergie du signal filtré sur la bande 8-13 Hz (significative de l'éveil), le quatrième est filtré sur la bande 4-7 Hz (significative du stade N1 et du sommeil paradoxal, c'est-à-dire le stade R) et sur le dernier graphique, le signal est filtré sur la bande 0,5-2 Hz (significative du sommeil profond, c'est-à-dire du stade N3). En rouge, nous avons fait figurer les correspondances entre l'énergie du signal filtré et le ou les stades de sommeil dont la bande de fréquence est caractéristique.

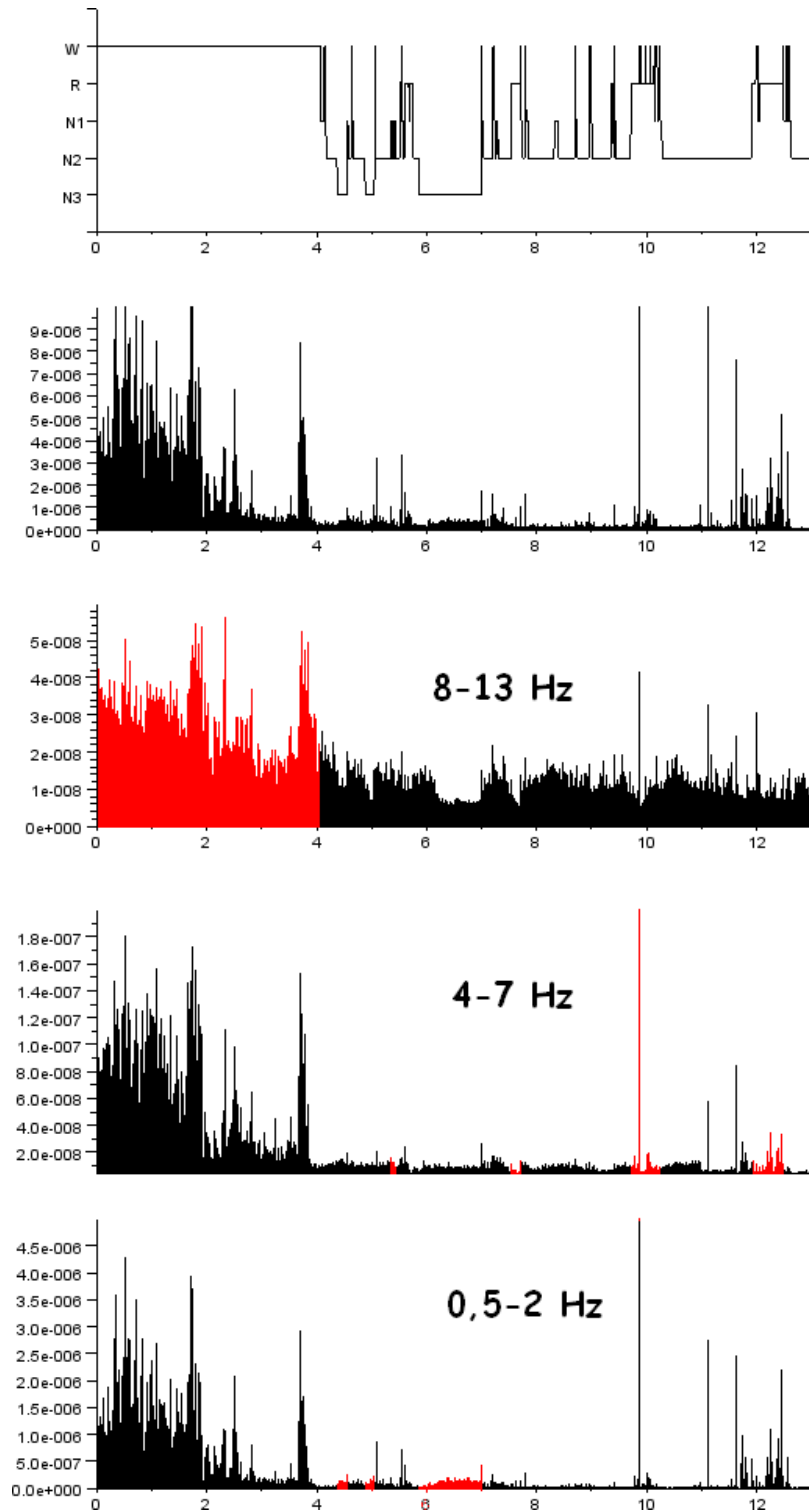


FIGURE 2.7 – Énergie par stade de sommeil et bandes de fréquences en fonction du temps

Nous voyons clairement que la distinction entre les différents stades de sommeil en utilisant la décomposition de Fourier et l'énergie du signal n'est pas un problème simple à résoudre avec les seuils définis par l'AASM. C'est pourquoi nous avons choisi d'utiliser un apprentissage supervisé avec les données toutes interprétées par un expert médical.

## 2.4 Choix des attributs d'entrée

Au total, 24 attributs sont utilisés : 8 par type de voie (EEG, EOG et EMG).

Ces 8 paramètres sont répartis en deux parties :

- les paramètres « globaux », c'est-à-dire calculés sur l'intégralité de l'époque ;
- les paramètres « locaux », c'est-à-dire calculés sur des fenêtres d'une seconde.

Pour les courbes EEG et EOG, plusieurs voies sont utilisées. Il est donc nécessaire d'établir des règles pour agréger les scores des différentes courbes. La règle générale est de calculer chaque paramètre pour chacune des voies enregistrées. Le score agrégé est la moyenne des paramètres sur l'ensemble des voies de même type. Nous verrons que pour certains paramètres, les règles d'agrégation appliquées respectent des cas particuliers.

### – Paramètres globaux,

#### 1. puissance du signal ;

Le premier paramètre est la puissance du signal, défini ainsi :  
pour un signal réel  $x$ , sa puissance vaut :

$$Puissance(x) = \frac{1}{T \cdot f_s} \times \sum (x(t))^2$$

où :

- $x(t)$  est l'amplitude du signal  $x$  à l'instant  $t$
- $T$  est la durée du signal  $x$  (30 secondes dans notre cas)
- $f_s$  est la fréquence d'échantillonnage (*sampling rate* ou *sampling frequency*)

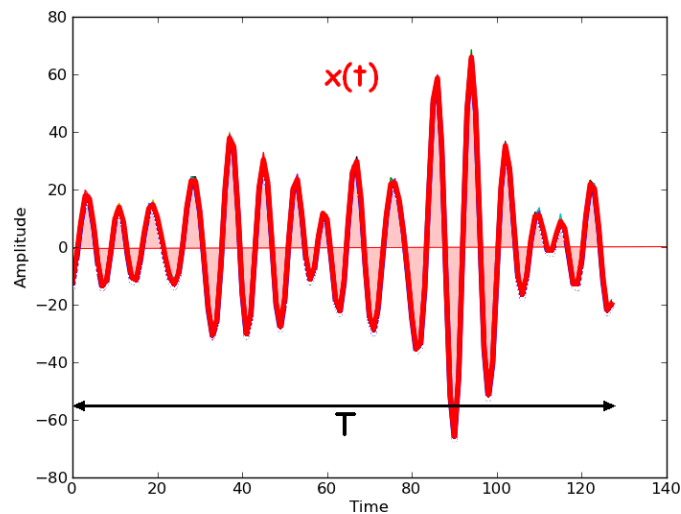


FIGURE 2.8 – Puissance du signal

## 2. Principale classe de fréquence ;

Le deuxième paramètre est la classe principale de fréquence.

Il est calculé à partir des résultats d'une transformée de Fourier Rapide (**F**ast **F**ourier **T**ransform ou *FFT*). La FFT est un algorithme de calcul de la Transformée de Fourier Discrète<sup>1</sup> permettant de passer du domaine temporel au domaine fréquentiel.

Les coefficients de Fourier sont regroupés par classes de fréquences définies dans le tableau 2.1. Les classes de fréquences EEG sont reconnues et utilisées par les experts médicaux, mais nous avons défini nous-mêmes les seuils des classes de fréquences pour les courbes EOG et EMG. À chaque classe est associée la médiane de tous les coefficients de Fourier correspondant à des fréquences comprises entre les bornes de la classe. La valeur maximum des coefficients agrégés résultants correspond à la classe utilisée comme deuxième paramètre.

Lorsqu'il y a plusieurs courbes de même type, le coefficient est calculé pour chaque courbe. On calcule ensuite la moyenne entre les 3 identifiants (numériques) de classe afin d'établir la valeur du paramètre.

TABLEAU 2.1 – Classes de fréquences EEG, EOG and EMG

class identifier	class label	EEG	EOG	EMG
1	$\delta$	0-4 Hz	0-2 Hz	0-2 Hz
2	$\theta$	4-7 Hz	2-3 Hz	2-3 Hz
3	$\alpha$	8-12 Hz	3-20 Hz	3-4 Hz
4	$\beta$	12-30 Hz	$> 20Hz$	$> 4Hz$

## 3. Fréquence du coefficient de Fourier maximum ;

Le troisième paramètre est la valeur de la fréquence à laquelle est associé le coefficient de Fourier maximum.

## 4. Classe de fréquences la plus fréquente localement ;

Le quatrième paramètre correspond à la classe de fréquences la plus fréquente parmi les  $N \times 30$  fenêtres locales d'une seconde qui composent l'époque, N étant le nombre de courbes de même type (EEG, EOG, EMG). Sur chacune de ces époques est calculée la classe de fréquences la plus fréquente (comme pour le paramètre 2). La classe la plus fréquente sur les  $N \times 30$  fenêtres étudiées donne le quatrième paramètre.

---

1. Soient  $x(t)$ ,  $t=0, \dots, n-1$  les  $n$  enregistrements successifs d'un signal. La formule de la  $f_k^{me}$  composante de la Transformée de Fourier Discrète de cet enregistrement est la suivante :

$$S(f_k) = \sum_{j=0}^{n-1} \left( x(j) \times e^{-2\pi i \frac{jk}{n}} \right), k = 0, \dots, n-1$$



## – Paramètres locaux,

5. Proportion de fenêtres locales d’une seconde de classe de fréquences principale  $\beta$  ;

Le cinquième paramètre correspond à la proportion de fenêtres locales d’une seconde pour lesquelles on a obtenu comme classe de fréquences principale la classe  $\beta$ .

6. Proportion de fenêtres locales d’une seconde de classe de fréquences principale  $\alpha$  ;

Le sixième paramètre correspond à la proportion de fenêtres locales d’une seconde pour lesquelles on a obtenu comme classe de fréquences principale la classe  $\alpha$ .

7. Proportion de fenêtres locales d’une seconde de classe de fréquences principale  $\theta$  ;

Le septième paramètre correspond à la proportion de fenêtres locales d’une seconde pour lesquelles on a obtenu comme classe de fréquences principale la classe  $\theta$ .

8. Proportion de fenêtres locales d’une seconde de classe de fréquences principale  $\delta$  ;

Le huitième paramètre correspond à la proportion de fenêtres locales d’une seconde pour lesquelles on a obtenu comme classe de fréquences principale la classe  $\delta$ .

## 2.5 Élaboration de l’arbre

Pour construire l’arbre, nous avons utilisé la plateforme WEKA (cf Figure 2.9) Il s’agit d’un logiciel rassemblant de nombreux algorithmes d’apprentissage automatique (*machine learning*) couramment utilisés pour élaborer un classifieur et ajuster ses paramètres. Cet outil est particulièrement simple à configurer et permet de tester plusieurs algorithmes sur les mêmes ensembles d’apprentissage et/ou de test.

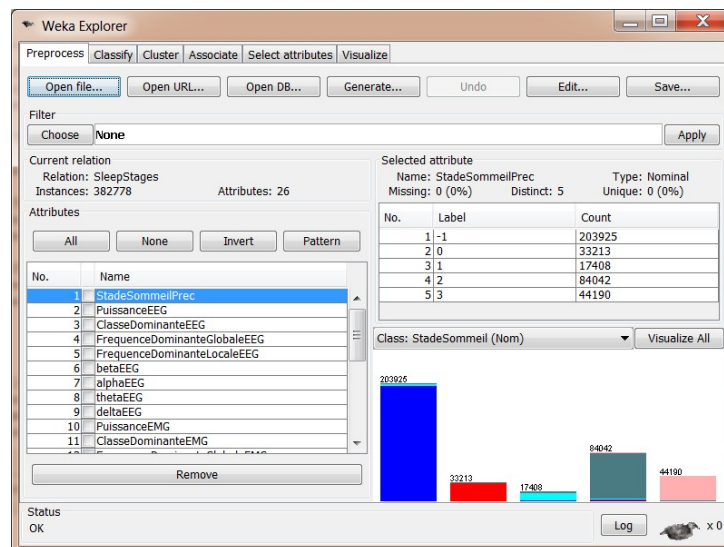


FIGURE 2.9 – Plateforme WEKA

## 2.6 Résultats

Nous avons testé différents algorithmes d'apprentissage pour construire un arbre de décision. WEKA a été configuré pour utiliser 33% des données pour l'apprentissage et le reste pour les tests. Chaque algorithme a été testé avec ses options par défaut. Afin d'améliorer les résultats, le métaclassifieur Bagging, appelé également Bootstrap aggregating, a été utilisé. Ce métaclassifieur construit, à partir de l'ensemble d'apprentissage, et par tirages aléatoires successifs avec remise, plusieurs sous-ensembles d'apprentissage, afin de générer plusieurs modèles et ainsi, déterminer lequel permet d'obtenir les meilleurs résultats.

Les résultats de ces tests sont donnés dans le tableau 2.2. Nous donnons deux statistiques :

- le pourcentage d'instances correctement classifiées ;

$$A = \frac{N_{OK}}{N_0}$$

où  $N_{OK}$  est le nombre d'instances correctement classifiées et  $N_0$  le nombre total d'instances dans l'ensemble de test.

- le coefficient Kappa de Cohen [15], qui tient compte de la taille des classes et donc de la probabilité de classer 2 instances dans la même classe de façon aléatoire.

$$\kappa = \frac{A - Pr(e)}{1 - Pr(e)}$$

où

- $A$  est le pourcentage d'instances correctement classifiées ;
- $Pr(e)$  est la probabilité d'un accord aléatoire.

$$Pr(e) = \frac{1}{N^2} \times \sum_{i=1}^r Q_i^1 \times Q_i^2$$

- $N$  est le nombre d'instances dans l'ensemble de tests ;
- $r$  est le nombre de classes du classifieur ;
- $Q_i^1$  et  $Q_i^2$  sont les nombres d'instances classées dans la classe  $i$  par, respectivement, le premier et le deuxième modèles.

TABLEAU 2.2 – Résultats des tests sur WEKA

Classifieur	Correctly Classified Instances	Kappa Statistic
<b>REPTree</b>	<b>78,0692%</b>	<b>0,6512</b>
J48	77,8933%	0,6499
SimpleCart	75,8431%	0,616
J48Graft	73,1608%	0,5781
NBTree	70,5027 %	0,5569
LADTree	71,474 %	0,5333
RandomTree	69,2944 %	0,5252
DecisionStump	61,2896%	0,3864

Le coefficient Kappa de Cohen est une mesure quantitative de l'accord. JR LANDIS et GG KOCH ont défini dans [35] une interprétation qualitative de Kappa en fonction de sa valeur, comme indiqué dans le tableau 2.3.

TABEAU 2.3 – Interprétation du coefficient Kappa de Cohen

$\kappa$	Interprétation
0,8 - 1,0	Accord presque parfait
0,6 - 0,8	Accord fort
0,4 - 0,6	Accord modéré
0,2 - 0,4	Accord faible
0,0 - 0,2	Accord très faible
<0	Désaccord

Compte tenu de ce tableau et des résultats obtenus sur l'ensemble de test, nous constatons que le meilleur modèle, c'est-à-dire l'arbre de décision construit par l'algorithme REPTree, donne des résultats avec un accord fort avec l'expert. De plus, ce modèle obtient un taux d'accord avec l'expert qui est de l'ordre de grandeur des accords entre plusieurs experts humains [17,18,22,57,69,72], qui se situe entre 70% et 90%.

Néanmoins, il convient d'analyser les résultats par stade de sommeil.

Pour cela, nous allons analyser dans un premier temps le *rappel*, la *précision* (ou *sensibilité*) et la *spécificité* pour chacune des classes de notre modèle avec l'arbre de décision obtenu par l'algorithme d'apprentissage REPTree, c'est-à-dire les différents stades de sommeil. Ces résultats sont donnés dans le tableau 2.4.

Le *rappel* est une mesure utilisée dans la communauté scientifique, de la capacité d'un système de décision à trouver le maximum d'instances d'une classe. Pour une classe donnée, elle mesure le taux d'instances correctement identifiées comme appartenant à cette classe.

La *précision* est une mesure utilisée dans la communauté scientifique, de la capacité d'un système de décision à être pertinent dans sa classification. Pour une classe donnée, elle mesure le taux d'instances considérées comme appartenant à cette classe qui appartiennent effectivement à cette classe. Dans la communauté médicale, elle correspond à la *sensibilité*.

La *spécificité* est une mesure utilisée dans la communauté médicale, de la capacité d'un système de décision à détecter qu'une instance n'appartient pas à une classe donnée. Elle sert notamment à vérifier qu'un test ne va pas considérer comme malade un individu sain.

Pour une classe  $\mathcal{C}$  donnée, le *rappel* et la *précision* sont définis ainsi :

$$\begin{cases} \text{Rappel}_{\mathcal{C}} = \frac{VP}{VP+FP} \\ \text{Précision}_{\mathcal{C}} = \text{Sensibilité}_{\mathcal{C}} = \frac{VP}{VP+FN} \\ \text{Spécificité}_{\mathcal{C}} = \frac{VN}{VN+FP} \end{cases}$$

où, pour un problème de recherche d'instances appartenant à une classe  $\mathcal{C}$  (à l'aide d'un classifieur) :

- $VP$  est le nombre de *vrais positifs*, c'est-à-dire le nombre d'instances de la classe  $\mathcal{C}$  classées dans la classe  $\mathcal{C}$  par le classifieur.
- $FP$  est le nombre de *faux positifs*, c'est-à-dire le nombre d'instances n'appartenant pas à la classe  $\mathcal{C}$  classées dans la classe  $\mathcal{C}$  par le classifieur.
- $VN$  est le nombre de *vrais négatifs*, c'est-à-dire le nombre d'instances n'appartenant pas à la classe  $\mathcal{C}$  non classées dans la classe  $\mathcal{C}$  par le classifieur.
- $FN$  est le nombre de *faux négatifs*, c'est-à-dire le nombre d'instances de la classe  $\mathcal{C}$  non classées dans la classe  $\mathcal{C}$  par le classifieur.

TABLEAU 2.4 – Rappel, Précision et Spécificité

Stade de sommeil	Rappel	Précision / Sensibilité	Spécificité
W	92,3%	89,1%	87,2%
R	52,5%	64,4%	97,3%
N1	3,7%	38,2%	99,7%
N2	74,2%	61,1%	86,6%
N3	68,1%	74,1%	96,9%

Le rappel moyen est donc de 58,2%.

La précision (ou sensibilité) moyenne est de 65,4%.

La spécificité moyenne est de 93,5%.

Les résultats sont représentés sur le graphique 2.10.

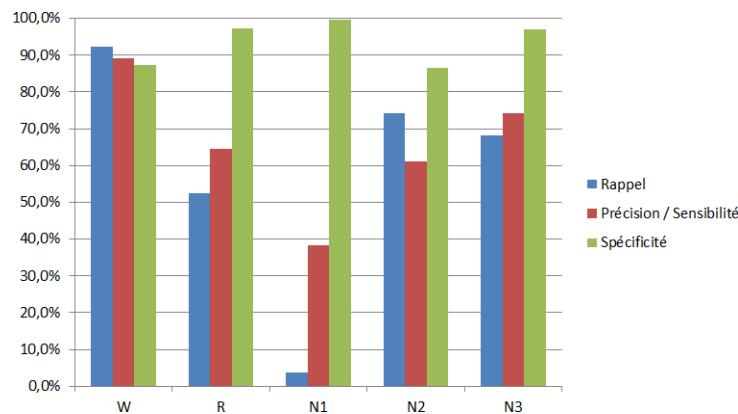


FIGURE 2.10 – Rappel et Précision

Pour analyser ces résultats de rappel et précision, nous allons maintenant calculer la F-mesure définie ainsi :

$$F = 2 \times \frac{\text{precision} \cdot \text{rappel}}{\text{precision} + \text{rappel}}$$

Ceci donne les résultats suivants, stade de sommeil par stade de sommeil, dans le tableau 2.5, représentés graphiquement sur la figure 2.11.

Stade de sommeil	F-mesure
W	90,7%
R	57,9%
N1	6,7%
N2	67,0%
N3	71,0%

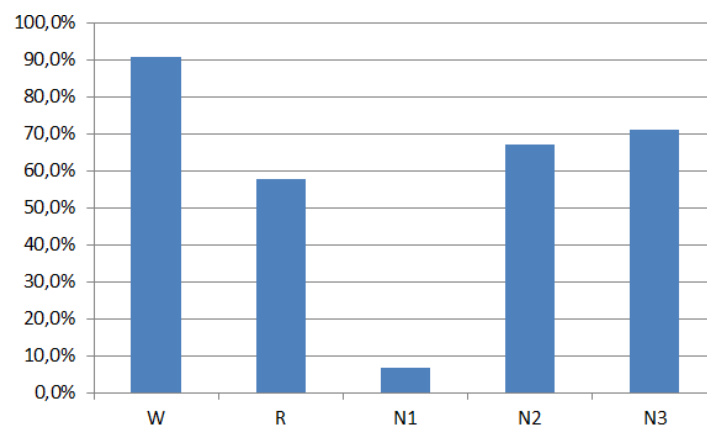


FIGURE 2.11 – F-Mesure

La F-Mesure globale vaut 61,6%.

Sous un regard purement statistique, avec un coefficient kappa valant 0,65 pour un taux d'époques correctement interprétées de 78%, l'algorithme semble bon en moyenne, mais avec une forte variabilité selon les stades de sommeil. En effet, l'algorithme distingue efficacement les périodes d'éveil aux périodes de sommeil et obtient des résultats corrects pour les stades N2 et N3, moyens pour le stade R et faibles pour le stade N1.

Il faut cependant nuancer la précision (ou sensibilité) obtenue pour l'éveil ; les enregistrements utilisés démarrent pour la plupart en fin d'après-midi, vers 17h ou 18h et durent 14 ou 15 heures. Au matin, l'infirmière remet au patient un questionnaire sur lequel il indique l'heure à laquelle il s'est couché et l'heure à laquelle il s'est réveillé. Cette période est utilisée pour l'analyse visuelle de l'expert. En dehors de cette période, il est considéré que le patient est éveillé, ce qui est vrai la majorité du temps, et rend donc les données exploitables pour l'apprentissage, mais entraîne un biais sur les résultats des tests en sous-estimant la précision réelle. En effet, l'algorithme peut détecter un assoupissement du patient avant son coucher, ce qui sera analysé comme une erreur.

Une des spécificités de notre ensemble de données est qu'il concerne des individus suspectés de souffrir d'un Syndrome d'Apnées du Sommeil et donc susceptible d'avoir un sommeil perturbé. Il est attendu que les tracés sont alors plus difficiles à interpréter, en particulier pour un algorithme de reconnaissance automatique. Nous avons donc

voulu analyser les résultats en fonction de la gravité de la pathologie chez les différents individus de notre base de données. Pour cela, nous avons implémenté sous scilab un parser de fichiers embl qui utilise le fichier de description de format EMBLA [51] et qui ensuite effectue l'extraction des paramètres époque par époque puis la classification avec l'arbre binaire de décision. Nous avons ensuite évalué les résultats sur les 71 patients dont nous avons les données polysomnographiques complètes (toutes les courbes valides sur toute la nuit) en les regroupant, au préalable, par sévérité de SAS. Cet ensemble de données est décrit dans le tableau 2.6.

Les résultats sont donnés dans le tableau 2.7 et illustrés sur la figure 2.12.

TABLEAU 2.6 – Description de l'ensemble des données utilisées pour les tests

sévérité	Nombre	Âge moyen	IAH moyen	% F % H
sain ( $IAH < 5$ )	24	56	2	33,33% 66,67%
léger ( $5 \leq IAH < 15$ )	24	48	11	25% 75%
modéré ( $15 \leq IAH < 30$ )	14	61	23	50% 50%
sévère ( $IAH \geq 30$ )	9	61	43	33,33% 66,67%

TABLEAU 2.7 – Taux de concordance de reconnaissance des stades de sommeil en fonction de la gravité du SAS

sévérité	N1	N2	N3	R
sain ( $IAH < 5$ )	19,98%	75,79%	71,77%	58,04%
léger ( $5 \leq IAH < 15$ )	19,62%	73,50%	68,78%	67,40%
modéré ( $15 \leq IAH < 30$ )	16,31%	68,07%	74,90%	52,43%
sévère ( $IAH \geq 30$ )	17,74%	63,00%	64,72%	53,94%

Nous observons la confirmation d'une dépendance de la qualité du classifieur automatique à la sévérité du SAS. En effet, plus l'IAH est élevé, et plus le taux de concordance est faible.

Pour une analyse plus en détails, il convient de comparer cet algorithme à d'autres algorithmes d'interprétation automatique des stades de sommeil présentés dans différentes publications, mais aussi aux résultats obtenus entre experts humains. En effet, il existe un désaccord entre experts humains sur l'interprétation des stades de sommeil, c'est pourquoi l'objectif d'un algorithme automatique est d'atteindre un niveau d'accord comparable à celui de deux experts humains entre eux, non seulement sur la globalité de l'interprétation mais aussi stade de sommeil par stade de sommeil.

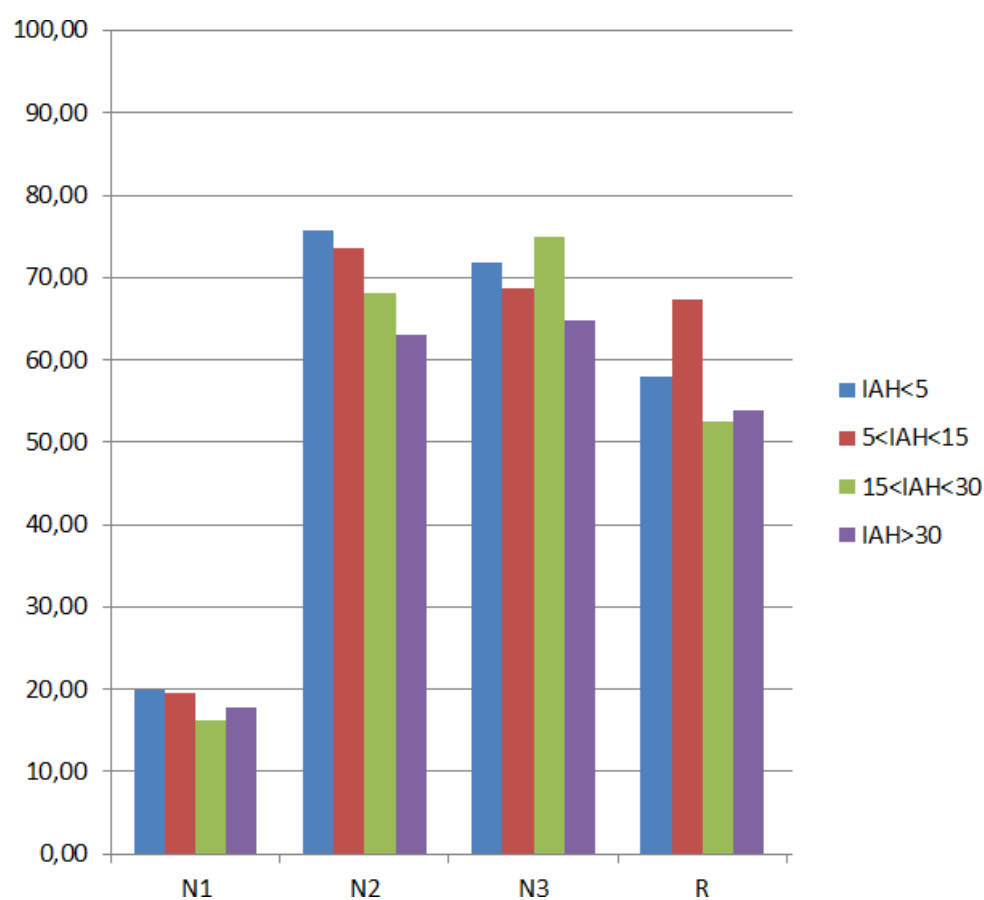


FIGURE 2.12 – Taux de concordance de reconnaissance des stades de sommeil en fonction de la gravité du SAS

## 2.7 Discussion

### Comparaison à d'autres algorithmes existants

La figure 2.13 illustre les résultats obtenus par les différents algorithmes de détection automatique des stades de sommeil regroupés par stade de sommeil. Nous pouvons constater que l'algorithme que nous proposons est bon pour détecter les périodes d'éveil, correct pour détecter le stade de sommeil N1 et le sommeil profond (N3), mais insuffisant pour détecter le stade N2 et le sommeil paradoxal (R).

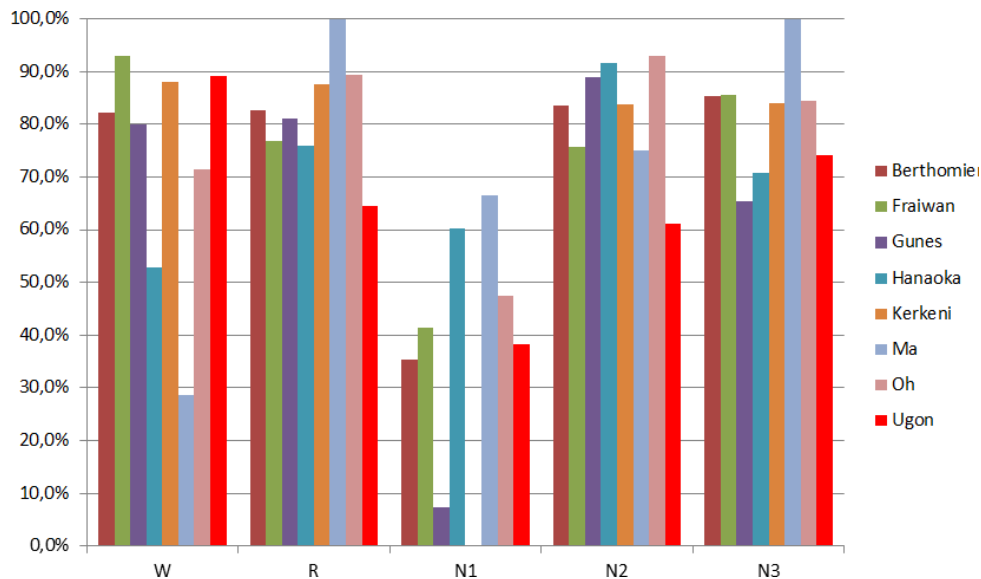


FIGURE 2.13 – Comparaison des algorithmes de reconnaissance de stades de sommeil

Cependant, le problème de la reconnaissance des stades de sommeil est un problème particulier pour lequel il existe des différences, parfois très importantes, d'un « scoreur » à l'autre. L'objectif d'un algorithme automatique est donc d'obtenir des résultats assimilables à ceux qu'un expert humain aurait. Un taux de concordance trop important témoigne donc d'un surapprentissage. Afin de pouvoir analyser les résultats des différents algorithmes, stade de sommeil par stade de sommeil, en prenant en compte les intervalles de taux d'accord inter- et intra-scoreurs issus des travaux de Michael H. SILBER dans [69], nous avons tracé la figure 2.14. Les résultats des différents algorithmes y sont représentés par des traits colorés. Les zones d'accord inter-scoreurs sont en jaune ; intra-scoreur en bleu. Les éventuels recouvrements des 2 zones sont en vert. L'objectif d'un algorithme est donc de se situer dans la zone jaune (ou verte) pour chacun des stades de sommeil.

Afin d'évaluer ces algorithmes, nous avons utilisé la grille de scores donnée dans le tableau 2.8, réalisée avec les conventions suivantes :

- Si, pour un stade de sommeil, le taux de concordance avec l'expert du classifieur est supérieur aux 2 bornes supérieures des concordances inter- et intra-scoreurs relevées par Michael H. SILBER dans [69], on utilise la notation '++'. Il y est associé un score de 0/2. Ceci témoigne d'un surapprentissage.
- Si ce taux est supérieur à la borne supérieure des concordances inter-scoreurs, mais situé dans l'intervalle des concordances intra-scoreurs, on utilise la notation '+'. Il y est associé un score de 1/2.



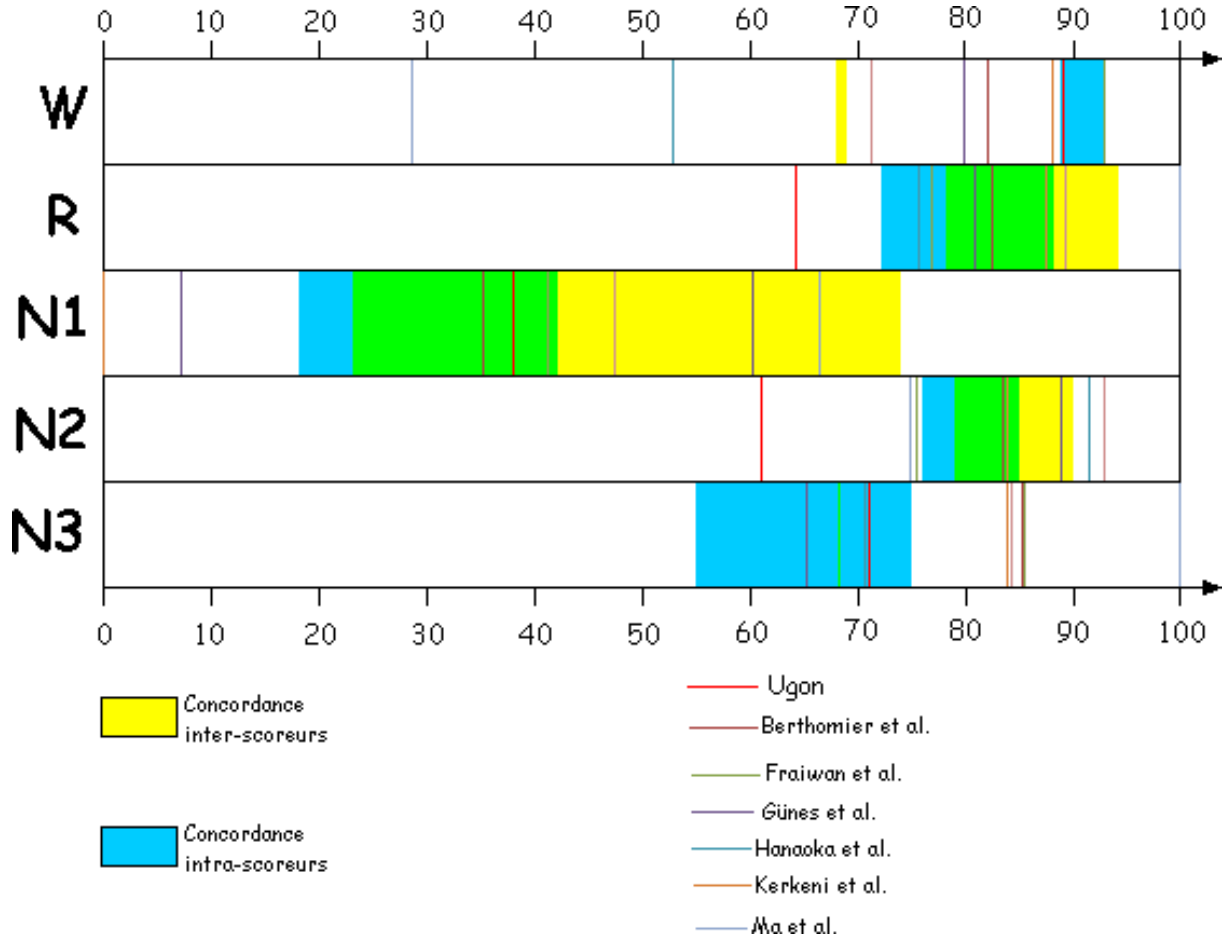


FIGURE 2.14 – Comparaison des algorithmes de reconnaissance de stades de sommeil avec prise en compte des intervalles d'accord entre scoreurs humains

- Si ce taux est compris dans l'intervalle des concordances inter-scoreurs, on utilise la notation 'OK'. Il y est associé un score de 2/2.
- Si ce taux est inférieur à la borne inférieure des concordances inter-scoreurs, mais situé dans l'intervalle des concordances intra-scoreurs, on utilise la notation '-'. Il y est associé un score de 1/2.
- si ce taux est inférieur aux 2 bornes inférieures des concordances inter- et intra-scoreurs, on utilise la notation '- -'. Il y est associé un score de 0/2.

L'analyse de ce tableau suggère que notre algorithme se situe dans la moyenne des algorithmes proposés dans la littérature ces dernières années. Néanmoins, une lecture attentive des conditions de réalisation des tests permet de nuancer ce classement moyen :

- L'évaluation faite dans la publication de Christian BERTHOMIER [10] repose sur les analyses visuelles effectuées indépendamment par 2 scoreurs humains et ensuite comparées aux résultats de l'analyse automatique. Seules les époques sans artefact et avec un accord des 2 scoreurs humains sont conservées. Cette méthode permet de filtrer les époques les plus difficiles à interpréter et donc sujettes à des différences d'interprétation. Il est donc attendu un taux d'accord plus élevé qu'une analyse conservant l'ensemble des époques.

TABLEAU 2.8 – Évaluation des algorithmes de reconnaissance des stades de sommeil

Algorithme	W	R	N1	N2	N3	score global
Berthomier	+	OK	OK	OK	++	7
Fraiwan	+	-	OK	OK	++	6
Günes	+	OK	- -	OK	-	6
Oh	+	OK	OK	++	++	5
Ugon	+	- -	OK	- -	+	4
Hanaoka	- -	-	OK	++	+	4
Kerkeni	+	-	OK	OK	++	6
Ma	- -	++	OK	- -	++	2

- Dans les travaux présentés par Luay A. FRAIWAN [23], 32 enregistrements polysomnographiques de la base de données MIT-BIH database [64] sont utilisés. Ces enregistrements sont très incomplets : ils disposent de 4 à 7 courbes ; la durée totale de l’enregistrement varie de 1h17mn à 6h30mn. Peu d’informations sont fournies sur les données mises en ligne et on peut supposer qu’ici aussi, les époques ont été filtrées.
- Les travaux de Salih GÜNES [28] sont évalués sur 5 enregistrements polysomnographiques de patients sains d’une durée moyenne de 7 heures, avec une durée moyenne de sommeil de 6h15mn48s, sans information sur le nombre d’experts ayant interprété les tracés. Le sommeil d’individus sains est moins perturbé que celui d’une personne atteinte d’un SAS et peut donc être plus facile à analyser automatiquement. En particulier, les tracés EEG d’un patient atteints de SAS présentent de nombreux micro-éveils.
- Pour sa part, Jung-Su OH inclut dans ses données de test des patients atteints de SAS. À nouveau, on ne dispose pas d’informations sur le nombre d’experts ayant interprété les enregistrements et sur une éventuelle sélection des époques utilisées.
- Les travaux de Masaaki Hanaoka [30] utilisent comme base de tests un enregistrement polysomnographique du sommeil d’un homme de 28 ans sain ; les courbes ont été analysées et validées par un collège d’experts. Les tests montrent un surapprentissage des stades N2 et N3 qui représentent près de 75% du temps d’un sommeil standard, ce qui permet d’avoir un taux de concordance élevé pas forcément représentatif d’un algorithme efficace . Il est notamment probable qu’il y ait eu un surapprentissage spécifique des caractéristiques de sommeil liées à cet individu.
- Nizar KERKENI n’a également utilisé, dans ses travaux [32], pour son apprentissage et ses tests que l’enregistrement polysomnographique d’un seul individu, ce qui cause probablement un surapprentissage du cas de cet individu.

### 3 Reconnaissance des stades de sommeil par fusion de paramètres symboliques

Dans cette deuxième méthode, nous allons proposer un algorithme basé sur la fusion symbolique. Cette méthode va consister, dans un premier temps, à extraire, à partir des courbes polysomnographiques, des informations interprétées sémantiquement et, dans un deuxième temps, à fusionner ces informations par une logique de prédicats. Afin de prendre en compte les courbes de même type (EEG, EOG), nous aurons besoin d’une

étape intermédiaire pour fusionner des paramètres symboliques de premier niveau pour générer des paramètres symboliques de second niveau.  
Ce travail a été effectué en collaboration avec Mlle Xue LIU.

### 3.1 Extraction de paramètres symboliques

Nous allons définir l'ensemble des paramètres symboliques qui vont être utilisés pour la fusion. Chaque paramètre symbolique est défini par un paramètre numérique dont la valeur est interprétée selon plusieurs seuils. Dans les travaux présentés ici, la valeur de ces seuils a été définie manuellement pour chaque paramètre indépendamment les uns des autres.

#### Paramètres symboliques de premier niveau

##### ***EEGLowWavesEnergy1* - Énergie des ondes lentes**

Ce premier paramètre mesure l'énergie des composantes du signal EEG situées sur la bande de fréquences 0-2 Hz.

Pour un signal EEG donné, la méthode pour déterminer ce paramètre est fournie dans le schéma bloc de la figure 2.15.

Il nécessite le calcul de la Transformée de Fourier Rapide (**F**ast **F**ourier **T**ransform ou FFT). Soient  $x_0, x_1, \dots, x_N$  les  $N$  points du signal  $x$ , la  $j$ -ème composante de la FFT du signal  $x$  vaut :

$$f_j = \sum_{k=0}^{N-1} x_k \times e^{-\frac{2i\pi}{N} \cdot j \cdot k}, j = 0, \dots, N-1$$

Il est également nécessaire de calculer l'énergie d'un signal discret  $x_0, x_1, \dots, x_N$  selon la formule suivante :

$$E = \sum_{k=0}^N |x_k|$$

Les valeurs de l'indicateur  $\vartheta$  sont interprétées selon les règles suivantes :

- Si  $\vartheta > 0,6$  alors *EEGLowWavesEnergy1* = « High ».
- Si  $0,42 < \vartheta \leq 0,6$  alors *EEGLowWavesEnergy1* = « Middle ».
- Si  $\vartheta \leq 0,42$  alors *EEGLowWavesEnergy1* = « Low ».

##### ***EEGProportionLowInstFreq* - Proportion de fréquences instantanées inférieures à 2 Hz**

Ce deuxième paramètre va calculer, en chaque point du signal EEG, la fréquence instantanée. Ensuite, pour chaque époque, nous allons mesurer le taux de fréquence instantanée inférieure à 2 Hz.

La méthode de calcul de ce paramètre nécessite l'utilisation d'une fenêtre de Hamming. L'équation de la courbe constituant la « fenêtre de Hamming » de  $N$  points est la suivante :

$$\omega(k) = 0,54 - 0,46 \times \cos\left(\frac{2 \cdot \pi \cdot k}{N-1}\right), 0 \leq k < N-1$$

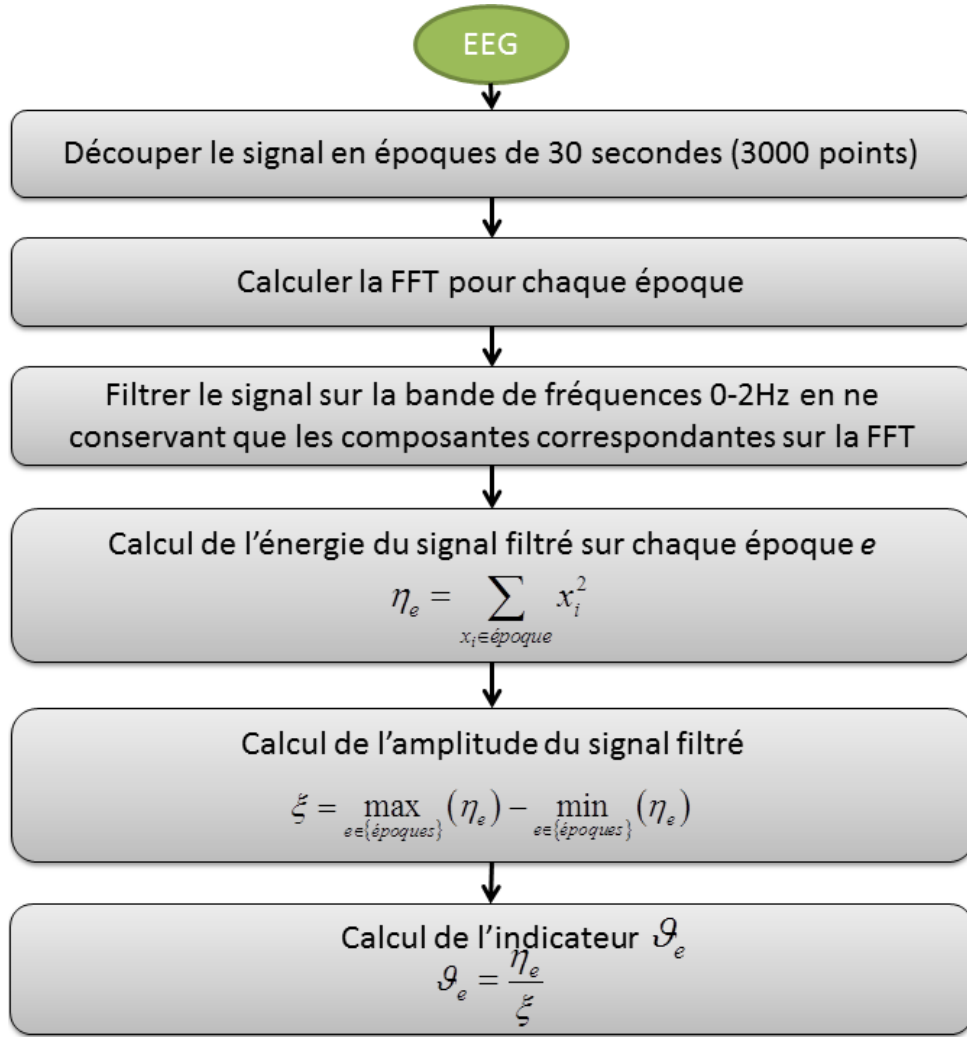


FIGURE 2.15 – Schéma bloc du paramètre « Énergie des ondes lentes »

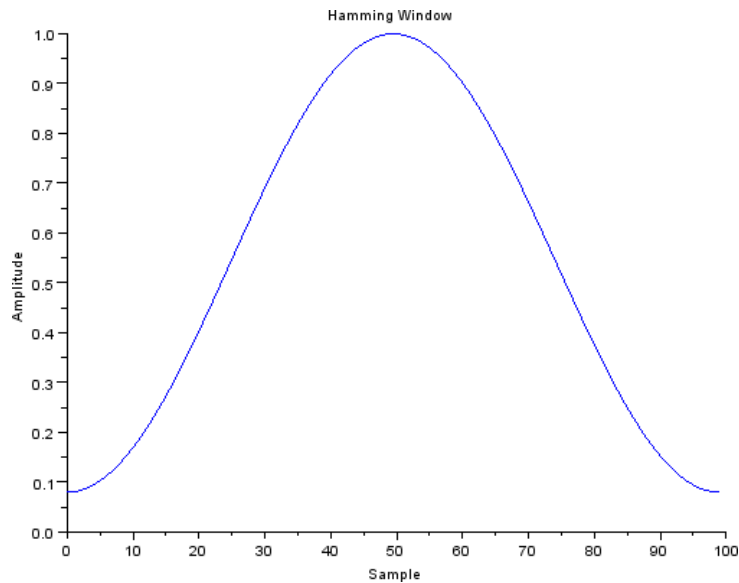


FIGURE 2.16 – Fenêtre de Hamming

Pour un signal EEG donné, la méthode pour déterminer ce paramètre est fournie dans le schéma bloc de la figure 2.17.

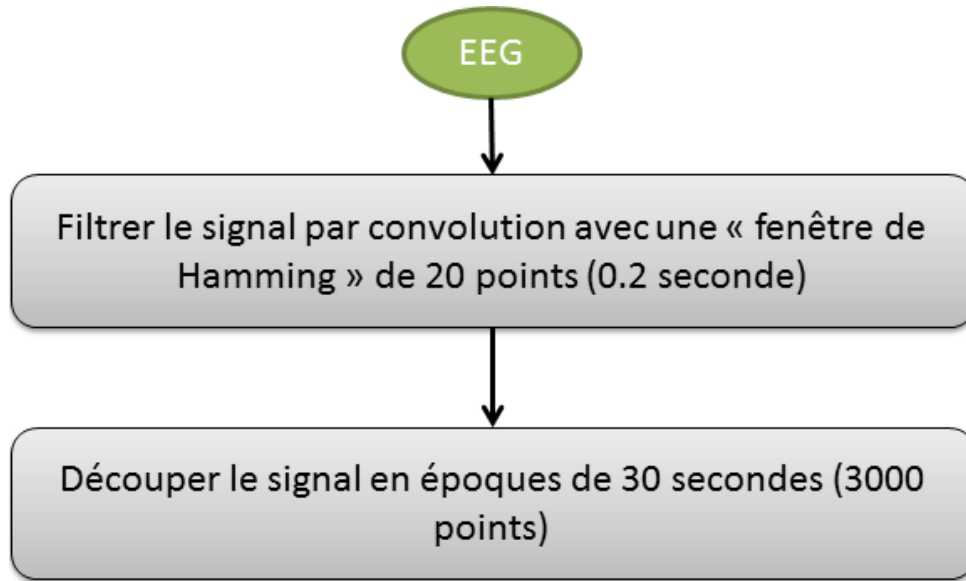


FIGURE 2.17 – Schéma bloc du calcul de la Fréquence Instantanée

On calcule ensuite, pour chaque époque,  $\rho_{0-2}$ , le taux de Fréquences Instantanées inférieures à 2 Hz.

Les valeurs de  $\rho_{0-2}$  sont ensuite interprétées selon les règles suivantes :

- Si  $\rho_{0-2} > 62\%$  alors  $EEGProportionLowInstFreq = \text{« High »}$ .
- Si  $45\% < \rho_{0-2} \leq 62\%$  alors  $EEGProportionLowInstFreq = \text{« Middle »}$ .
- Si  $\rho_{0-2} \leq 45\%$  alors  $EEGProportionLowInstFreq = \text{« Low »}$ .

#### ***EEGProportionThetaInstFreq* - Proportion de fréquences instantanées entre 4 et 7 Hz**

Ce troisième paramètre est très similaire au précédent. Il utilise également le calcul de la fréquence instantanée et mesure sa proportion de valeurs situées entre 4 et 7 Hz, intervalle de fréquences des ondes  $\theta$  pour les ondes cérébrales.

Le calcul de la fréquence instantanée reprend la méthode du schéma bloc 2.17.

On calcule ensuite, pour chaque époque,  $\rho_{4-7}$ , le taux de Fréquences Instantanées situées entre 4 et 7 Hz.

Les valeurs de  $\rho_{4-7}$  sont ensuite interprétées selon les règles suivantes :

- Si  $\rho_{4-7} > 80\%$  alors  $EEGProportionThetaInstFreq = \text{« High »}$ .
- Si  $\rho_{4-7} \leq 80\%$  alors  $EEGProportionThetaInstFreq = \text{« Low »}$ .

#### ***EEGVariabilite1* - Variabilité du signal EEG**

La variabilité du signal EEG est un paramètre qui indique si le signal EEG varie beaucoup dans le temps, par une mesure indirecte du taux d'ondes rapides (supérieures à 18 Hz) particulièrement élevé. Ces ondes rapides sont caractéristiques de l'éveil. Afin d'éviter la multiplicité des calculs, nous allons reprendre les résultats de la Fréquence Instantanée, obtenue par la méthode présentée dans le schéma bloc 2.17.

On calcule ensuite, pour chaque époque,  $\rho_{18+}$ , le taux de Fréquences Instantanées supérieures à 18 Hz.

Les valeurs de  $\rho_{18+}$  sont ensuite interprétées selon les règles suivantes :

- Si  $\rho_{18+} > 12\%$  alors  $EEGVariabilite1 = \text{« Unstable »}$ .
- Si  $5\% < \rho_{18+} \leq 12\%$  alors  $EEGVariabilite1 = \text{« Middle »}$ .
- Si  $\rho_{18+} \leq 5\%$  alors  $EEGVariabilite1 = \text{« Stable »}$ .

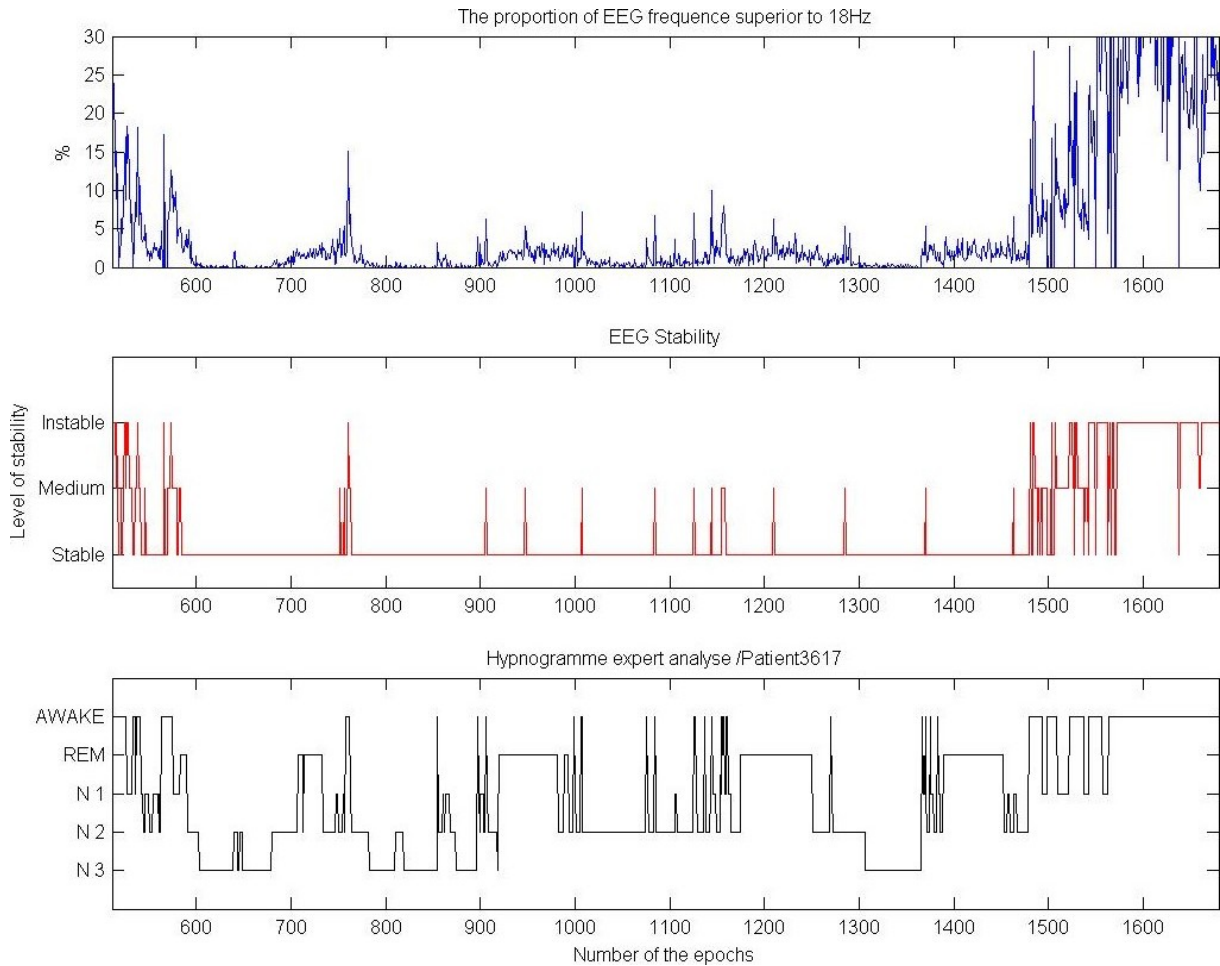


FIGURE 2.18 – Comparaison entre les valeurs du paramètre  $EEGVariabilite1$  et l'hypnogramme

La figure 2.18 met en relation les valeurs de  $\rho_{18+}$ , de  $EEGVariabilite1$  avec l'hypnogramme d'un patient. On peut y voir une corrélation entre le stade *éveil* et la valeur « Unstable » du paramètre  $EEGVariabilite1$ .

### ***EEGSleepSpindles1* - Présence de fuseaux du sommeil**

Ce paramètre permet de savoir s'il y a présence de fuseaux du sommeil (Sleep Spindles) sur une voie EEG donnée lors d'une époque donnée. Il nécessite donc la reconnaissance des fuseaux du sommeil sur un signal EEG. D'après l'AASM [31], les fuseaux du sommeil sont définis comme *un train d'ondes distinctes de fréquence entre 11 et 16 Hz avec une durée supérieure à 0,5 seconde*.

La méthode de reconnaissance des fuseaux du sommeil va nécessiter l'utilisation des éléments suivants :

- Le filtre à réponse impulsionnelle finie (**F**inite **I**mpulse **R**esponse ou FIR) est un filtre numérique linéaire dont la sortie dépend uniquement de l'entrée. Il s'oppose au filtre à réponse impulsionnelle infinie dont la sortie peut dépendre à la fois des valeurs d'entrée mais aussi des valeurs passées de la réponse. Généralement, il s'écrit comme une convolution du signal sur une durée finie (de taille N points) avec une fenêtre prédéfinie de taille N.
- La fonction Gamma, d'équation :

$$\Gamma(n) = (n - 1)!, \text{ pour } n > 0$$

- La fonction de Bessel modifiée d'ordre  $\nu$ , d'équation :

$$i_\nu(z) = \left(\frac{z}{2}\right)^\nu \times \sum_{k=0}^{\infty} \frac{\left(\frac{z^2}{4}\right)^k}{k! \cdot \Gamma(\nu + k + 1)}$$

qui est un ensemble de fonctions solutions de l'équation différentielle :

$$x^2 \cdot \frac{d^2 y}{dx^2} + x \cdot \frac{dy}{dx} + (x^2 - \nu^2) \cdot y = 0$$

- La fenêtre de Kaiser de paramètre  $\beta$ , dont l'équation, pour une fenêtre de N points, est la suivante :

$$\omega(k) = \frac{I_0\left(\beta \times \sqrt{1 - \frac{4 \cdot k^2}{(N-1)^2}}\right)}{I_0(\beta)}, \text{ avec } -\frac{N-1}{2} \leq k \leq \frac{N-1}{2}$$

La figure 2.19 présente une illustration de la fenêtre de Kaiser avec  $\beta = 6$ .

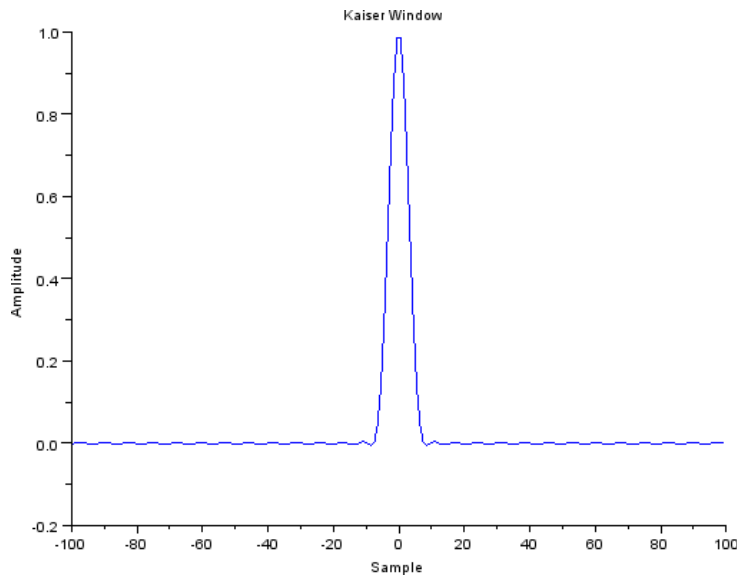


FIGURE 2.19 – Fenêtre de Kaiser avec  $\beta = 6$

- La Transformée de Fourier à court terme (**S**hort **T**ime **F**ourier **T**ransform ou STFT) calcule la transformée de Fourier sur une portion locale du signal grâce à une fenêtre glissante. Le signal est considéré comme nul en dehors de la fenêtre glissante. On peut alors appliquer la Transformée de Fourier Rapide « Traditionnelle ».
- La « fenêtre de Hamming » de N points dont l'équation est la suivante :

$$\omega(k) = 0,54 - 0,46 \times \cos\left(\frac{2 \cdot \pi \cdot k}{N-1}\right), 0 \leq k < N-1$$

Une illustration de la fenêtre de Hamming est donnée sur la figure 2.16.

- Le spectrogramme est le module au carré de la Transformée de Fourier à court terme (STFT). Il correspond au niveau d'énergie de chaque composante sur chaque segment de signal analysé par la STFT.

Pour un signal EEG donné, la méthode pour reconnaître et compter le nombre de fuseaux du sommeil est présenté dans le schéma bloc de la figure 2.20.

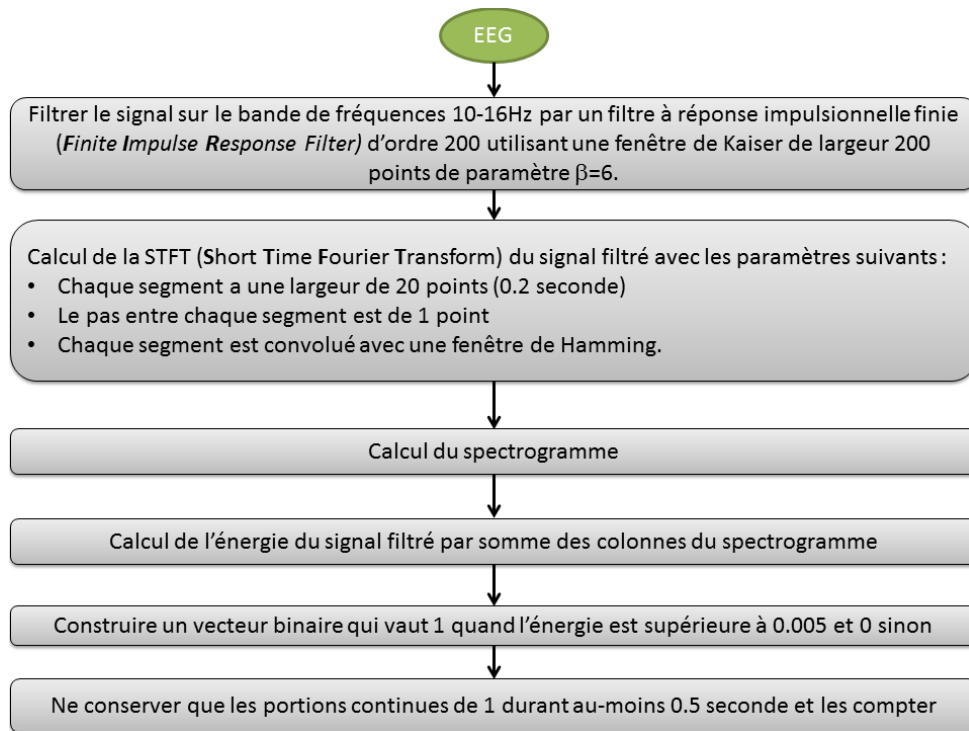


FIGURE 2.20 – Schéma bloc de l'algorithme de reconnaissance et dénombrement des fuseaux du sommeil

Afin d'illustrer les étapes de cette méthode, nous proposons en figure 2.21 5 graphiques. Le premier est le signal source sur une époque. Il s'agit de la dérivation EEG C4-A1. Le deuxième graphique montre le signal filtré par le filtre passe bande 10-16Hz. Le graphique suivant est le spectrogramme sur lequel apparaissent des tâches rouges indiquant des positions possibles de fuseaux du sommeil. L'énergie du signal est donnée sur le graphique suivant. Les zones sur fond rouge pâle entre barres verticales rouges indiquent les zones où le seuil de 0,005 est dépassé, sur une durée de plus de 0,5 secondes. Le résultat est fourni sur un graphique représentant la fréquence instantanée. Les fuseaux



du sommeil sont indiqués sur fond bleu clair. On peut y voir que la fréquence instantanée, bien qu'intéressante, serait difficile à interpréter car sa valeur ne reste pas entre 10 et 16 Hz sur toute la durée du fuseau du sommeil.

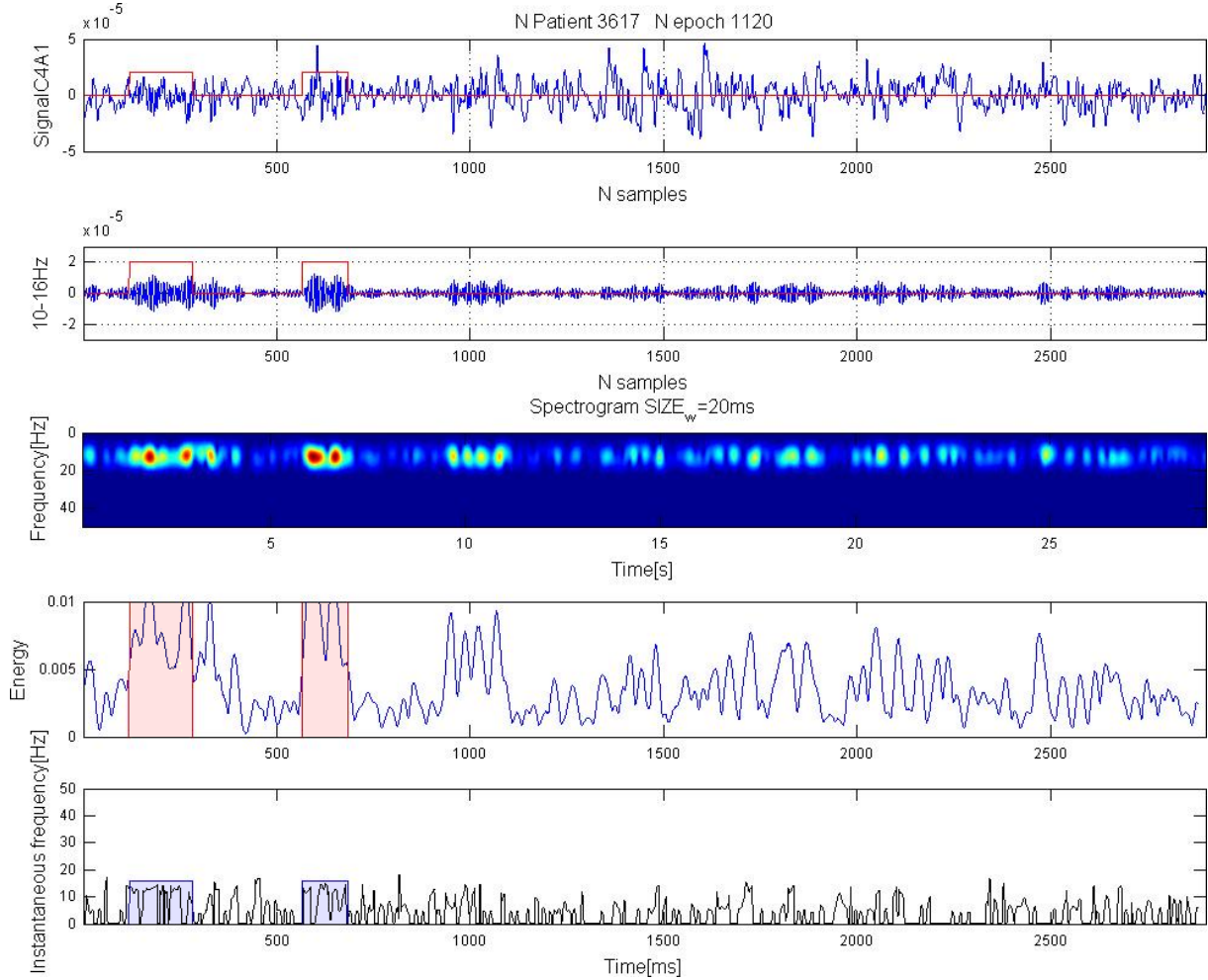


FIGURE 2.21 – Exemple réel de reconnaissance de fuseaux du sommeil

Les valeurs du nombre  $\sigma$  de fuseaux du sommeil sont ensuite interprétées selon les règles suivantes :

- Si  $\sigma \geq 1$  alors  $EEGSleepSpindles1 = \text{« Présence »}$ .
- Si  $\sigma = 0$  alors  $EEGSleepSpindles1 = \text{« Absence »}$ .

### ***EMGActivityLevel* - Niveau d'activité de l'EMG**

Ce paramètre est le paramètre d'analyse du signal EMG qui va indiquer le niveau d'activité du signal EMG, ce qui est une indication du tonus musculaire utilisé pour interpréter les stades de sommeil.

Le paramètre ***EMGActivityLevel*** correspond à la moyenne, pour une époque donnée, des valeurs des points du signal en valeur absolue normés pour chaque patient.

Pour un signal EMG donné, la méthode de calcul de ce paramètre est présentée dans le schéma bloc de la figure 2.22.

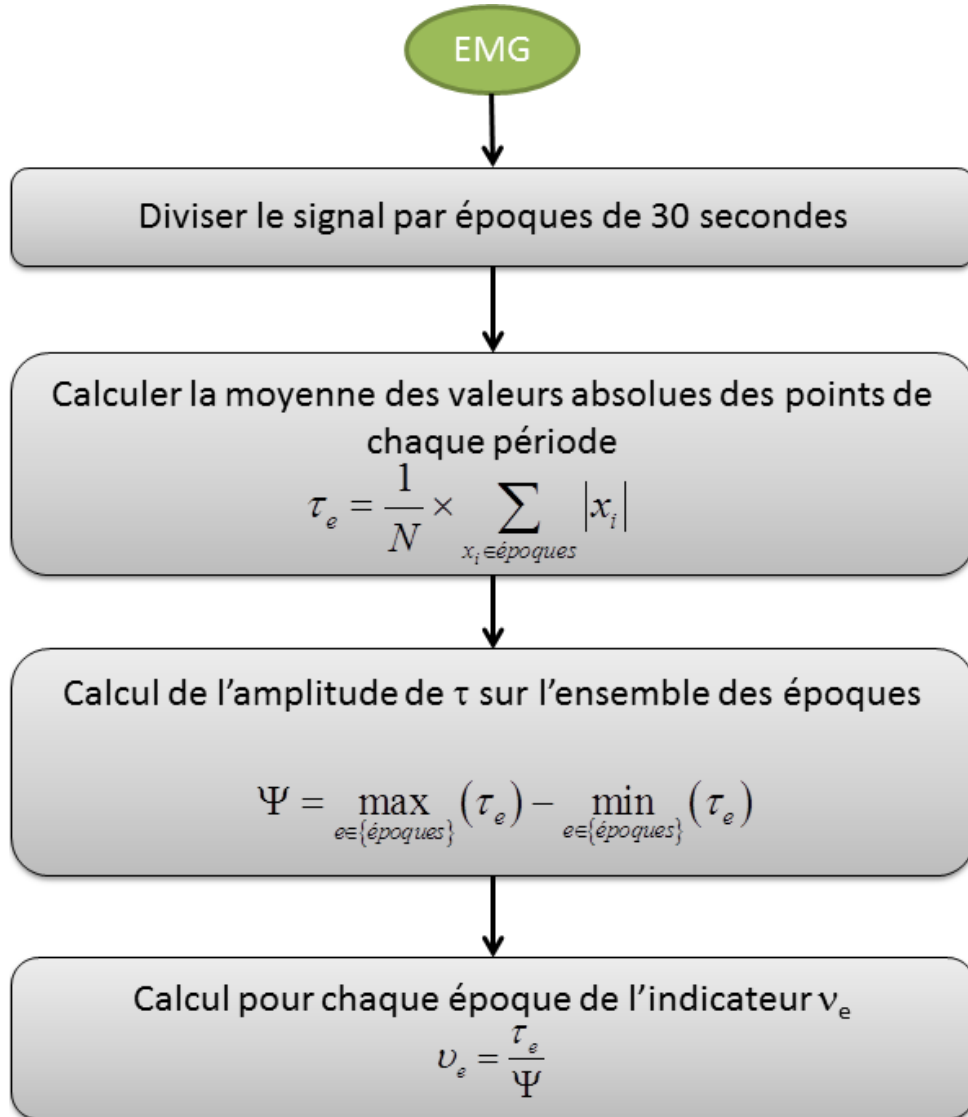


FIGURE 2.22 – Schéma bloc de l’algorithme de calcul de l’indicateur de mesure de l’activité EMG

Les valeurs de  $\nu_e$  sont ensuite interprétées selon les règles suivantes :

- Si  $\nu_e > 0,75$  alors *EMGActivityLevel*=« High ».
- Si  $0,5 < \nu_e \leq 0,75$  alors *EMGActivityLevel*=« Middle ».
- Si  $0,25 < \nu_e \leq 0,5$  alors *EMGActivityLevel*=« Low ».
- Si  $0 \leq \nu_e \leq 0,25$  alors *EMGActivityLevel*=« Lowest ».

La figure 2.23 compare les valeurs du paramètre *EMGActivityLevel* aux stades de sommeil sur l’hypnogramme pour 2 patients. Le premier graphique, en bleu, est une illustration du signal EMG ; sur ce graphique, à chaque époque est associée la valeur moyenne du signal *Submental EMG* sur cette époque. En dessous, en rouge, est fournie la valeur du paramètre pour chaque époque. Ces valeurs sont mises en relation avec l’hypnogramme tracé sur le troisième graphique.

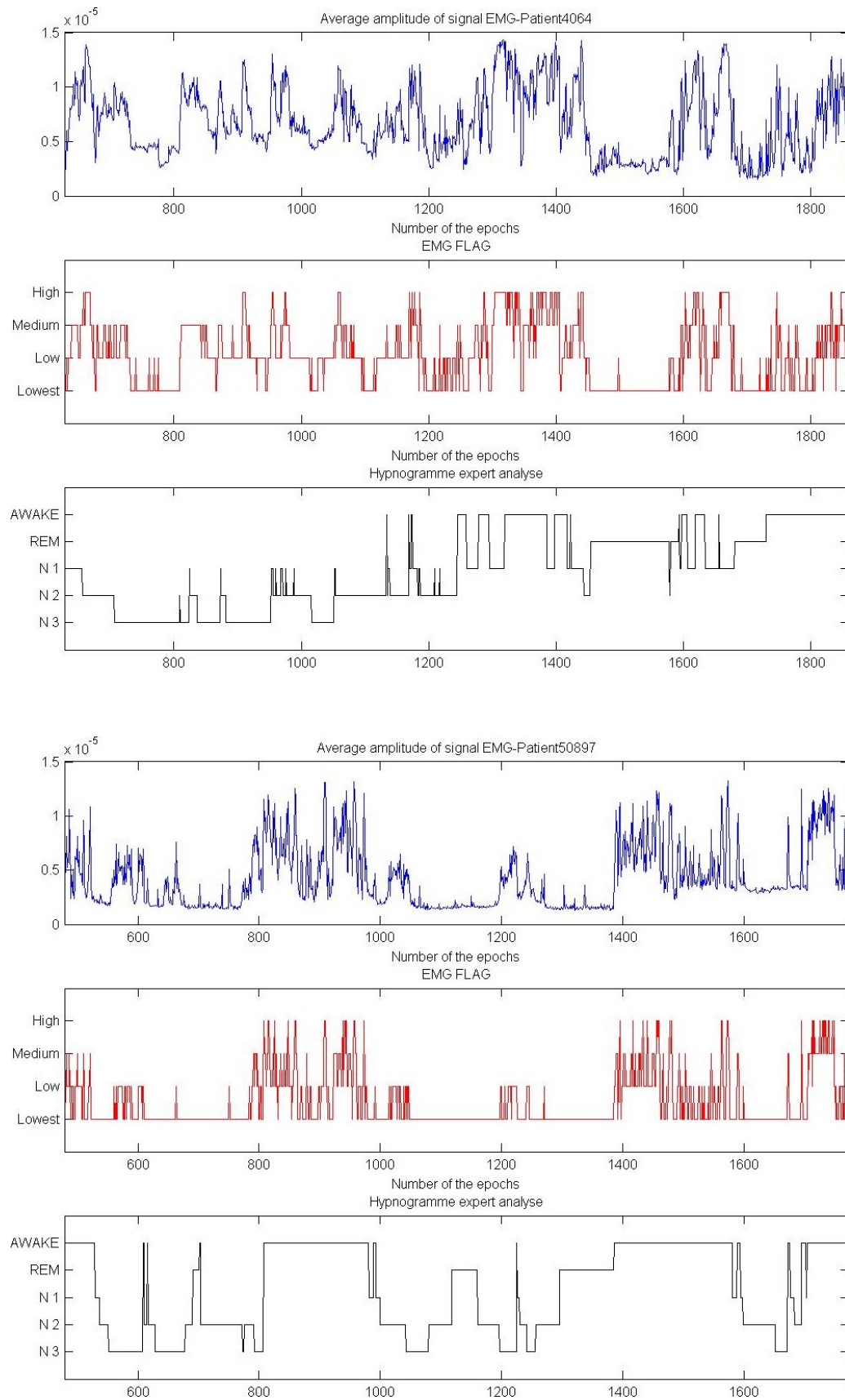


FIGURE 2.23 – Comparaison des valeurs du paramètre *EMGActivityLevel*

### *EOGMvtsOcul1* - Nombre de Mouvements oculaires

Ce paramètre est une estimation du nombre de mouvements oculaires. La méthode s'inspire des travaux de Manuel MERINO et al. dans [52].

Dans cette publication, Manuel MERINO *et al* propose une méthode de reconnaissance et classification des mouvements oculaires, avec pour objectif de pouvoir piloter un robot par les yeux. Un mouvement est donc composé d'un mouvement de l'iris partant du centre dans une direction suivi d'un second mouvement dans la même direction revenant vers le centre, comme indiqué sur la figure 2.24. Son dispositif nécessite donc 4 capteurs, pour différencier les mouvements horizontaux des mouvements verticaux. Ce dispositif est également capable d'identifier les clignements de paupières, pour lesquels le motif qui apparaît sur l'EOG a une forme identique à un mouvement oculaire, mais est beaucoup plus rapide.

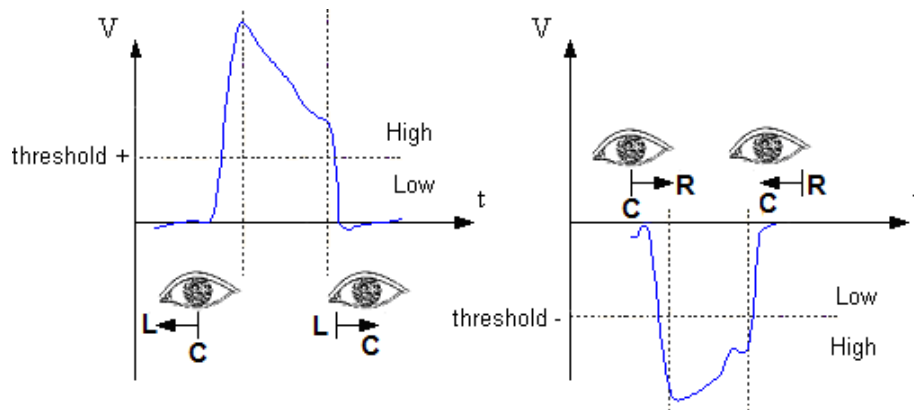


FIGURE 2.24 – Composantes des mouvements oculaires

L'application la reconnaissance des stades de sommeil ne nécessite pas de savoir distinguer la direction, ou le sens, de chaque mouvement oculaire, ni même les mouvements oculaires des clignements de paupières; seule leur quantification est utile à la reconnaissance des stades d'éveil ou de sommeil paradoxal. C'est donc une tâche beaucoup plus simple.

La détection des mouvements oculaires se fait par l'analyse de 3 données :

- la valeur du signal au point courant  $eog(t)$ ;
- la valeur du signal au point courant  $eog(t-n)$ ;
- la valeur de la dérivée.

$$eog'(t) = eog(t) - eog(t - n)$$

Il y a un mouvement oculaire si les conditions suivantes sont réunies :

- la courbe EOG forme un angle, c'est-à-dire que la dérivée a une valeur non nulle (voir Figure 2.25) ;
- la valeur de  $eog(t-n)$  est au niveau d'amplitude « *low* » (voir Figure 2.24) ;
- la valeur de  $eog(t)$  est au niveau d'amplitude « *high* » (voir Figure 2.24).

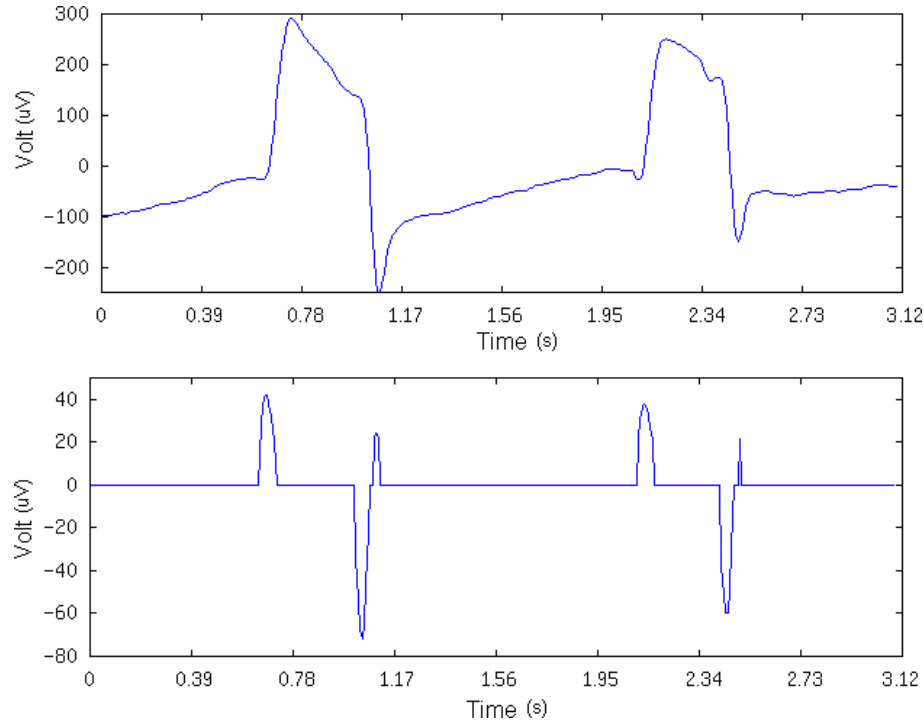


FIGURE 2.25 – Prétraitement du signal pour la reconnaissance des mouvements oculaires

Pour un signal EOG donné, la méthode de calcul de ce paramètre est présenté dans le schéma bloc de la figure 2.26.

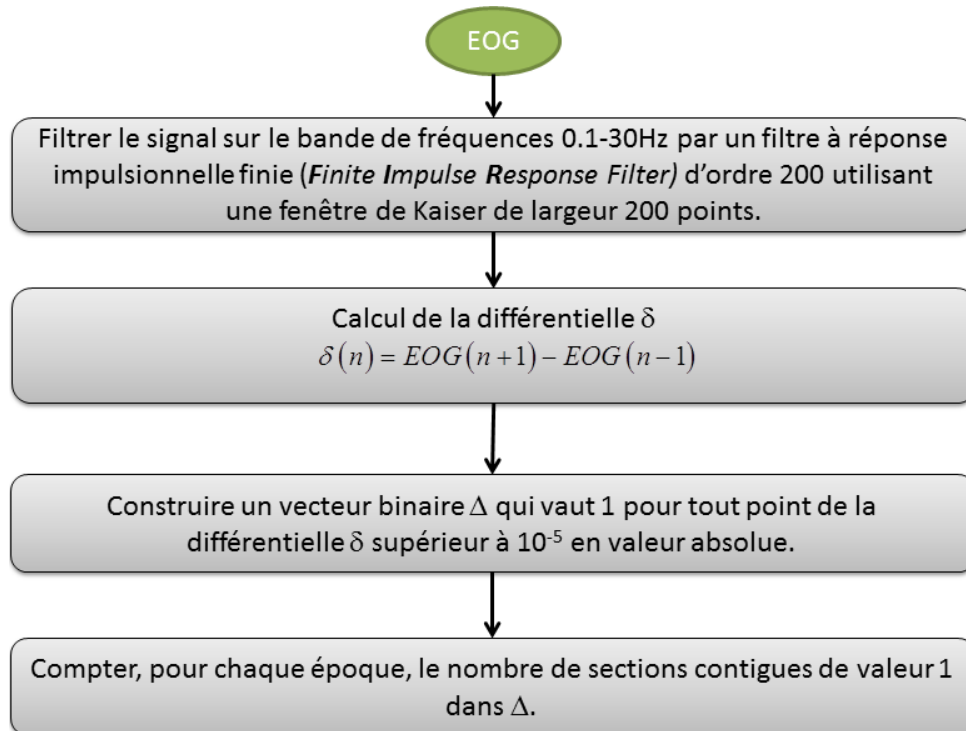


FIGURE 2.26 – Schéma bloc de l’algorithme de calcul de l’indicateur de quantité de mouvements oculaires  $EOGMvtsOcul1$

Soit  $\mu_e$  le nombre de mouvements oculaires trouvés sur une époque donnée. Les valeurs de  $\mu_e$  sont ensuite interprétées selon les règles suivantes :

- Si  $\mu_e \geq 20$  alors  $EOGMvtsOcul1 = \text{« high »}$ .
- Si  $10 < \mu_e < 20$  alors  $EOGMvtsOcul1 = \text{« Middle »}$ .
- Si  $2 < \mu_e \leq 10$  alors  $EOGMvtsOcul1 = \text{« Low »}$ .
- Si  $\mu_e \leq 2$  alors  $EOGMvtsOcul1 = \text{« Lowest »}$ .

### ***EOGCorr* - Corrélation Mouvements oculaires**

Ce dernier paramètre va mesurer le taux de corrélation entre l'EOG de l'œil droit et l'EOG de l'œil gauche, afin de voir si les mouvements sont conjugués, c'est-à-dire coordonnés, ou disconjugués, c'est-à-dire indépendants.

La formule pour calculer le taux de corrélation entre l'EOG de l'œil gauche (EOGL) et celui de l'œil droit (EOGR) est la suivante :

$$r = \frac{cov(EOGR, EOGL)}{\sigma_{EOGR} \times \sigma_{EOGL}} = \frac{\sum_{i=1}^n (EOGL - \overline{EOGL}) \cdot (EOGR - \overline{EOGR})}{\sqrt{\sum_{i=1}^N (EOGL - \overline{EOGL})^2} \times \sqrt{\sum_{i=1}^N (EOGR - \overline{EOGR})^2}}$$

où :

- $\overline{EOGL}$  est la moyenne des points EOGL sur l'époque.
- $\overline{EOGR}$  est la moyenne des points EOGR sur l'époque.

Les valeurs de  $r$  sont ensuite interprétées selon les règles suivantes :

- Si  $r \leq 0,4$  alors  $EOGCorr = \text{« Disconjugate »}$ .
- Si  $r > 0,4$  alors  $EOGNbMvts = \text{« Conjugate »}$ .

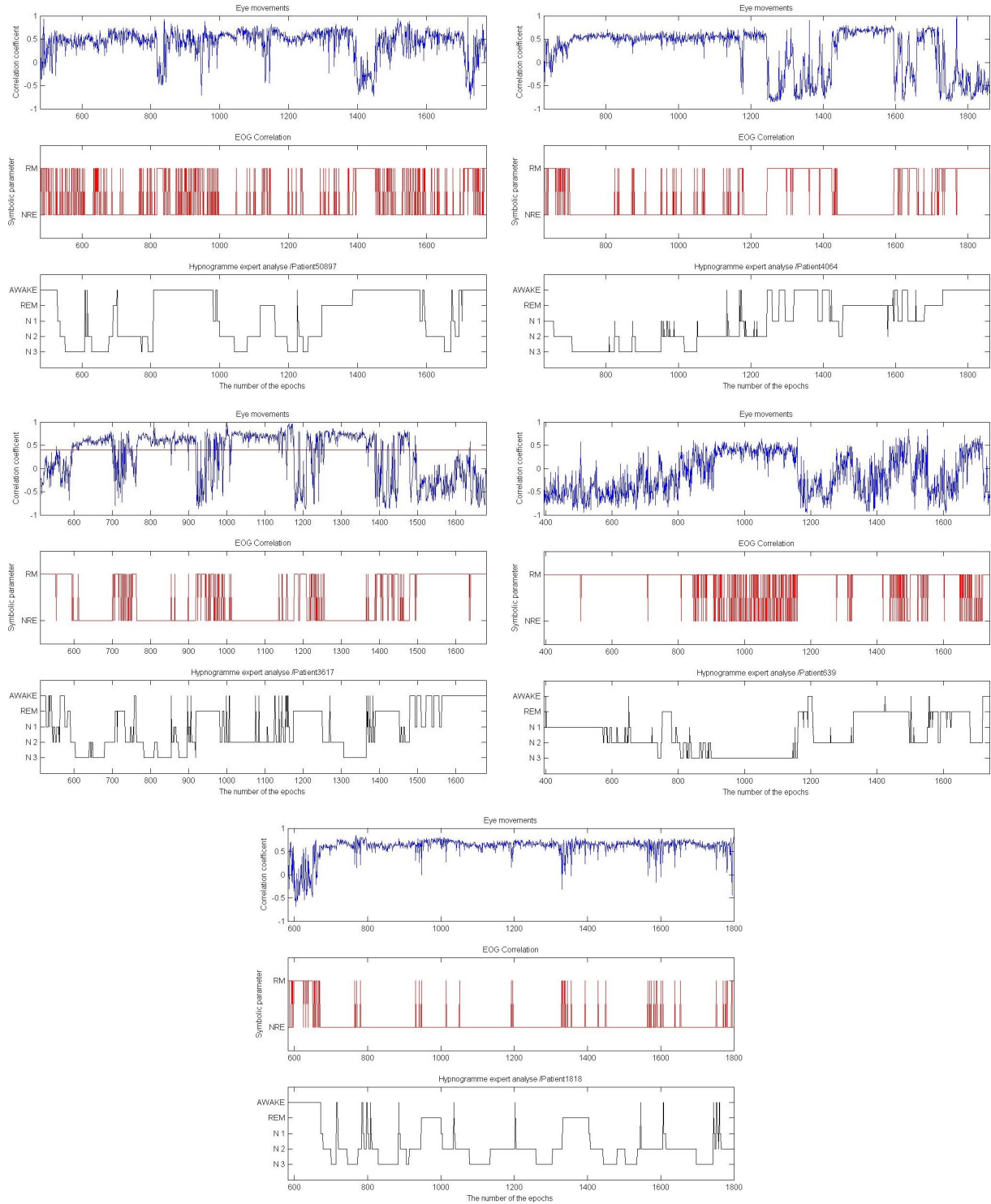


FIGURE 2.27 – Mise en évidence de l'utilité du paramètre  $EOG Corr$  pour reconnaître le sommeil paradoxal (REM)

Afin d'illustrer le lien entre la corrélation des mouvements oculaires et les stades de sommeil, nous fournissons sur la figure 2.27 cinq graphiques, pour 5 patients différents. Chaque graphique est constitué de trois courbes. La courbe du haut, en bleu, illustre le taux de corrélation entre les deux voies  $EOG-R$  et  $EOG-L$ . Le deuxième graphique donne les valeurs du paramètre  $EOG Corr$ . Ces deux graphiques sont mis en correspondance avec le troisième graphique qui est l'hypnogramme, construit à partir de l'interprétation visuelle de l'expert. On peut y voir une correspondance entre le sommeil paradoxal



(REM) et les mouvements oculaires disconjugués, c'est-à-dire avec une faible corrélation. On peut y voir aussi qu'en sommeil lent, les mouvements oculaires sont bien conjugués et qu'à l'éveil, les valeurs sont plus variables.

### Paramètres symboliques de second niveau

Les paramètres symboliques de second niveau sont utiles pour apporter de l'information au niveau d'un type de capteur donné. On les obtient en fusionnant les paramètres de premier niveau obtenus pour chaque dérivation pour un type de capteur donné.

#### ***EEGLowWavesEnergy2* - Énergie des ondes lentes**

Les valeurs de *EEGLowWavesEnergy2* sont obtenues par fusion des valeurs de *EEGLowWavesEnergy1* pour les 3 voies EEG : C4-A1, C3-A2, O1-A2. Sa valeur est obtenue en suivant les règles suivantes :

**Si** *EEGLowWavesEnergy1* = « High » pour au moins 2 dérivations parmi les 3 (C4-A1, C3-A2 et O1-A2)

**Alors** *EEGLowWavesEnergy2* = « High »

**Sinon** *EEGLowWavesEnergy2* = « Low »

**Fin Si**

La figure 2.28 met en correspondance les valeurs du paramètre *EEGLowWavesEnergy1* et celles du paramètre *EEGLowWavesEnergy2* avec l'hypnogramme. On peut y constater que la valeur « High » du paramètre *EEGLowWavesEnergy2* est associé au sommeil lent profond (N3).

#### ***EEGProportionLowInstFreq2* - Proportion de fréquence instantanée inférieure à 2 Hz**

Les valeurs de *EEGProportionLowInstFreq2* sont obtenues par fusion des valeurs de *EEGProportionLowInstFreq1* pour les 3 voies EEG : C4-A1, C3-A2, O1-A2. Sa valeur est obtenue en suivant les règles suivantes :

**Si** *EEGProportionLowInstFreq1* = « High » pour au moins 2 dérivations parmi les 3 (C4-A1, C3-A2 et O1-A2)

**Alors** *EEGProportionLowInstFreq2* = « High »

**Sinon** *EEGProportionLowInstFreq2* = « Low »

**Fin Si**



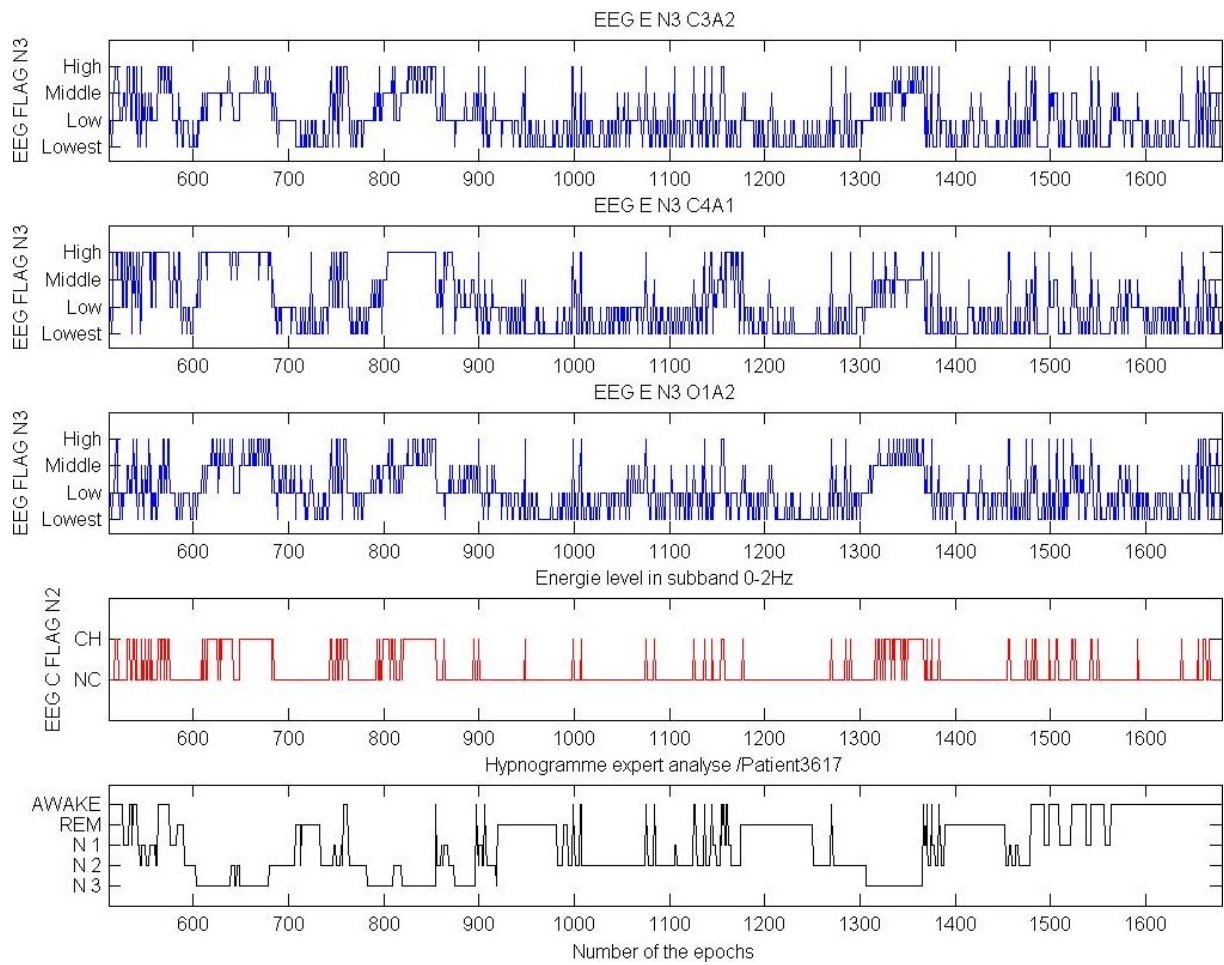


FIGURE 2.28 – Mise en évidence de l'utilité du paramètre *EEGLowWavesEnergy2* pour reconnaître les stades de sommeil

***EEGProportionThetaInstFreq2* - Proportion de fréquence instantanée entre 4 et 7 Hz**

Les valeurs de *EEGProportionThetaInstFreq2* sont obtenues par fusion des valeurs de *EEGProportionThetaInstFreq1* pour les 3 voies EEG : C4-A1, C3-A2, O1-A2. Sa valeur est obtenue en suivant les règles suivantes :

**Si** *EEGProportionThetaInstFreq1* = « High » pour au moins 2 dérivations parmi les 3 (C4-A1, C3-A2 et O1-A2)

**Alors** *EEGProportionThetaInstFreq2* = « High »

**Sinon** *EEGProportionThetaInstFreq2* = « Low »

**Fin Si**

***EEGVariabilite2* - Variabilité du signal EEG**

Les valeurs de *EEGVariabilite2* sont obtenues par fusion des valeurs de *EEGVariabilite1* pour les 3 voies EEG : C4-A1, C3-A2, O1-A2. Sa valeur est obtenue en suivant les règles suivantes :

**Si** *EEGVariabilite1* = « Unstable » pour au moins 1 des 3 dérivations (C4-A1, C3-A2 et O1-A2)

**Alors** *EEGVariabilite2* = « Unstable »

**Sinon, si** *EEGVariabilite1* = « Stable » pour les 3 dérivations

**Alors** *EEGVariabilite2* = « Stable »

**Sinon** *EEGVariabilite2* = « Unknown »

**Fin Si**

La figure 2.29 met en correspondance les valeurs du paramètre *EEGVariabilite1* et celles du paramètre *EEGVariabilite2* avec l'hypnogramme. On peut y constater que, dans le stade éveil, la valeur du paramètre vaut « Unstable ».

***EEGSleepSpindles2* - Présence de fuseaux du sommeil**

Les valeurs de *EEGSleepSpindles2* sont obtenues par fusion des valeurs de *EEGSleepSpindles1* pour les 3 voies EEG : C4-A1, C3-A2, O1-A2. Sa valeur est obtenue en suivant les règles suivantes :

**Si** *EEGSleepSpindles1* = « Présence » pour au moins 2 des 3 dérivations (C4-A1, C3-A2 et O1-A2)

**Alors** *EEGSleepSpindles2* = « Présence »

**Sinon** *EEGSleepSpindles2* = « Absence »

**Fin Si**

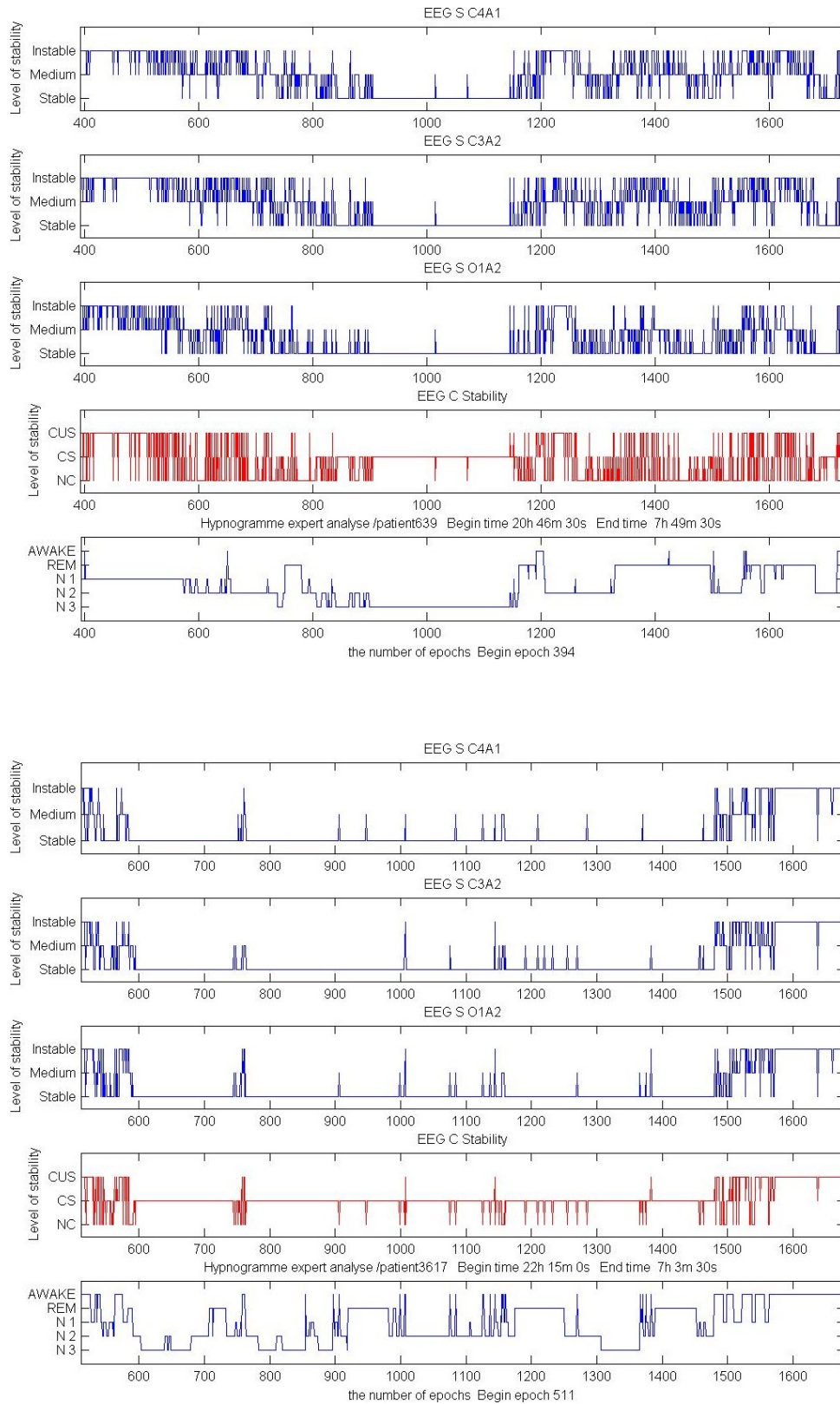


FIGURE 2.29 – Mise en évidence de l'utilité du paramètre *EEGVariabilite2* pour reconnaître les stades de sommeil

***EOGMvtsOcul2* - Nombre de Mouvements oculaires**

Les valeurs de *EOGMvtsOcul2* sont obtenues par fusion des valeurs de *EOGMvtsOcul1* pour EOG-L et EOG-R.

**Si** *EOGMvtsOcul1*(EOG-L)=« Lowest » **ET** *EOGMvtsOcul1*(EOG-R)=« Lowest » **Alors** *EOGMvtsOcul2*=« Lowest »

**Sinon,** **si** (*EOGMvtsOcul1*(EOG-L)=« Lowest » **OU** *EOGMvtsOcul1*(EOG-L)=« Low ») **ET** (*EOGMvtsOcul1*(EOG-R)=« Lowest » **OU** *EOGMvtsOcul1*(EOG-R)=« Low »)

**Alors** *EOGMvtsOcul2*=« Low »

**Sinon,** **si** *EOGMvtsOcul1*(EOG-L)=« High » **OU** *EOGMvtsOcul1*(EOG-R)=« High »

**Alors** *EOGMvtsOcul2*=« High »

**Sinon** *EOGMvtsOcul2*=« Middle »

**Fin Si**

Une illustration de ce paramètre est fournie en figure 2.30 pour deux patients. On peut y voir que la présence de mouvements oculaires est corrélée avec les stades W (éveil) et R (sommeil paradoxal).

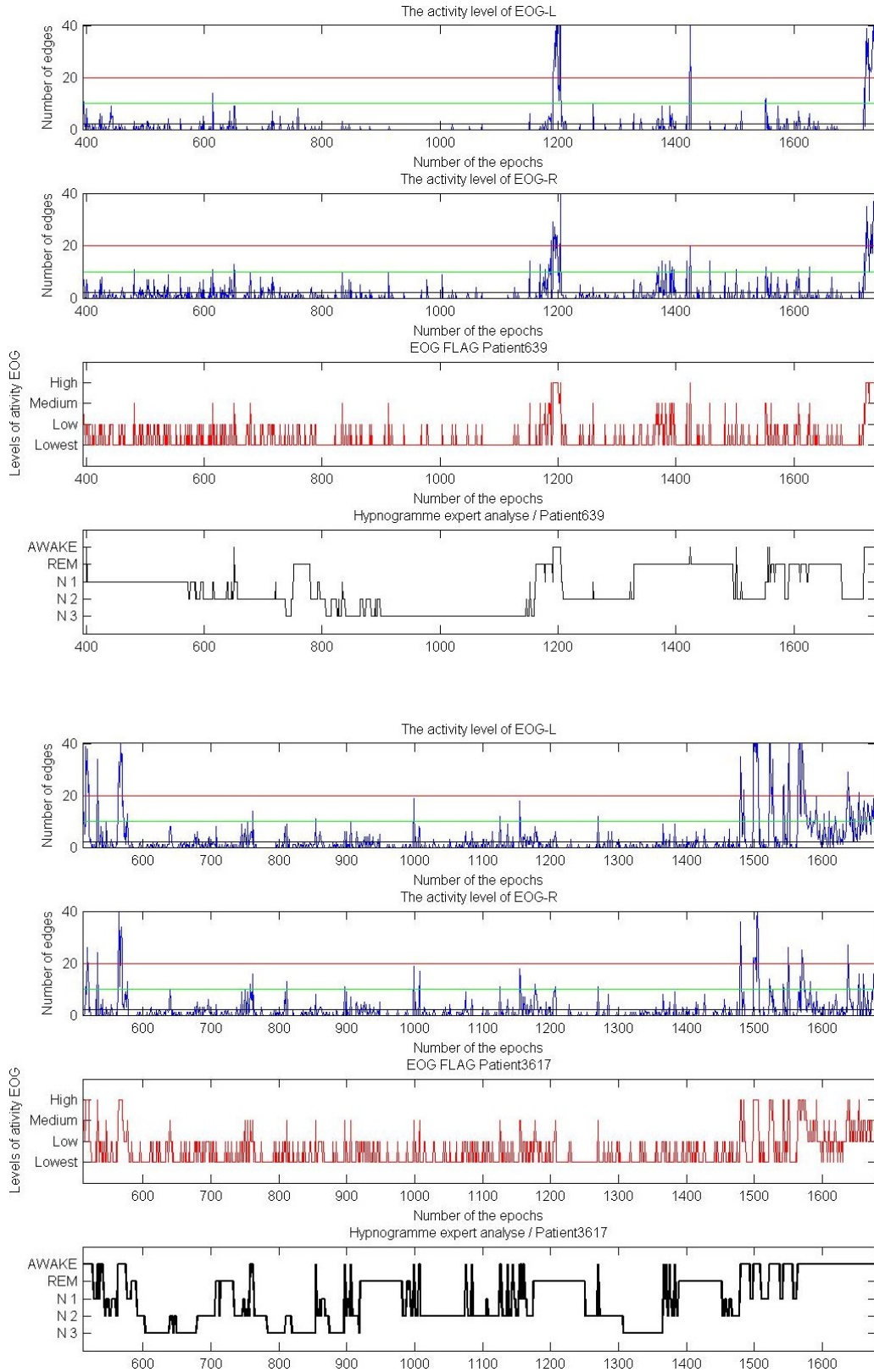


FIGURE 2.30 – Mise en relation du paramètre  $EOGMvtsOcul2$  avec l'hypnogramme

### 3.2 Identification des stades de sommeil par Fusion Symbolique

La reconnaissance des stades de sommeil se fait ensuite par Fusion Symbolique en utilisant les paramètres symboliques de second niveau. Les tests vont être effectués successivement, en plusieurs parcours des époques. Le premier passage permet d'identifier les époques d'éveil (stade W); le deuxième est effectué sur les époques restantes et permet d'identifier les époques de sommeil paradoxal (stade REM); les passages suivants permettent d'identifier, successivement, le stade N2, puis le stade N3. Les époques restantes sont considérées comme étant du stade N1.

Les règles utilisées pour la fusion symbolique vont être décrites sous la forme de diagrammes utilisant la symbolique décrite dans les figures 2.31, 2.32 et 2.33.

Sur la figure 2.31, la classe X est vraie si  $A = ValueA$  **OU**  $B = ValueB$  **OU**  $C = ValueC$ .

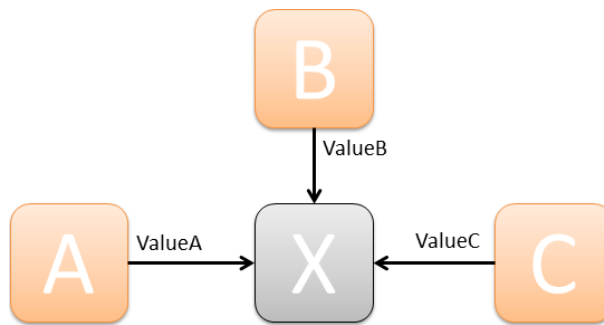


FIGURE 2.31 – Exemple de règle **OU**

Sur la figure 2.32, la classe Y est vraie si  $D = ValueD$  **OU**  $(E = ValueE$  **ET**  $F = ValueF)$ .

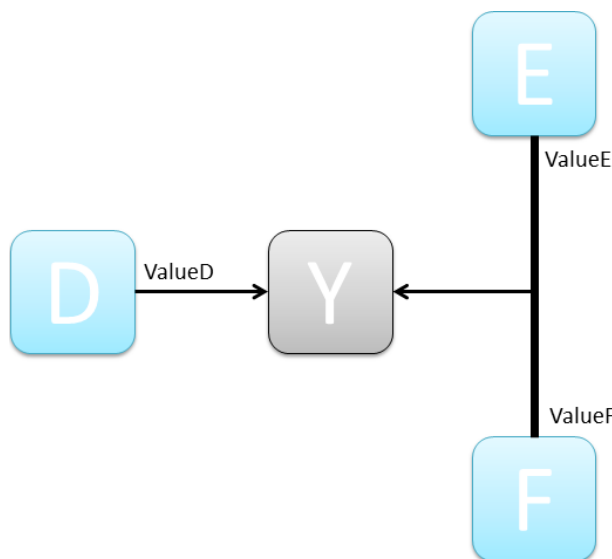


FIGURE 2.32 – Exemple de règle **ET**

Sur la figure 2.33, la classe Z est vraie si il existe au moins 2 variables qui valent *valueZ* parmi les variables G, H, I.

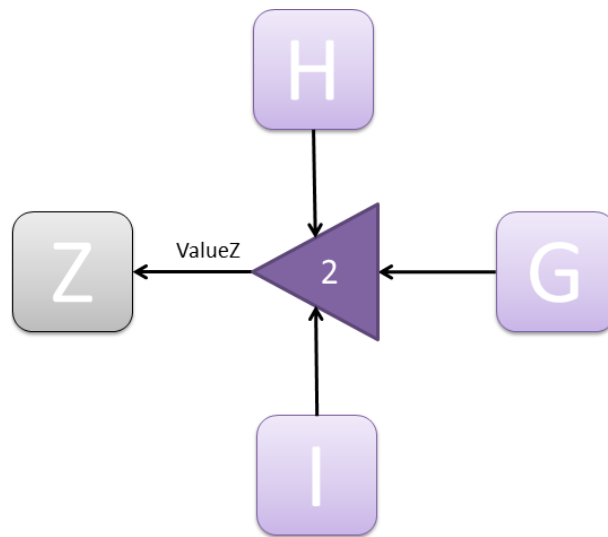


FIGURE 2.33 – Exemple de règle **AU MOINS**

La figure 2.34 illustre les règles pour tester si une époque est du stade W ou non.

Ainsi, une époque est scorée comme étant de stade W si l'une des 3 conditions suivantes est satisfaite :

- $EEGVariable2 = \text{« Unstable »}$ .
- $EOGMvtsOcul2 = \text{« High »}$ .
- $EOGMvtsOcul2 = \text{« Middle » ET } EMGActivityLevel \text{ IN « Middle », « High »}$ .

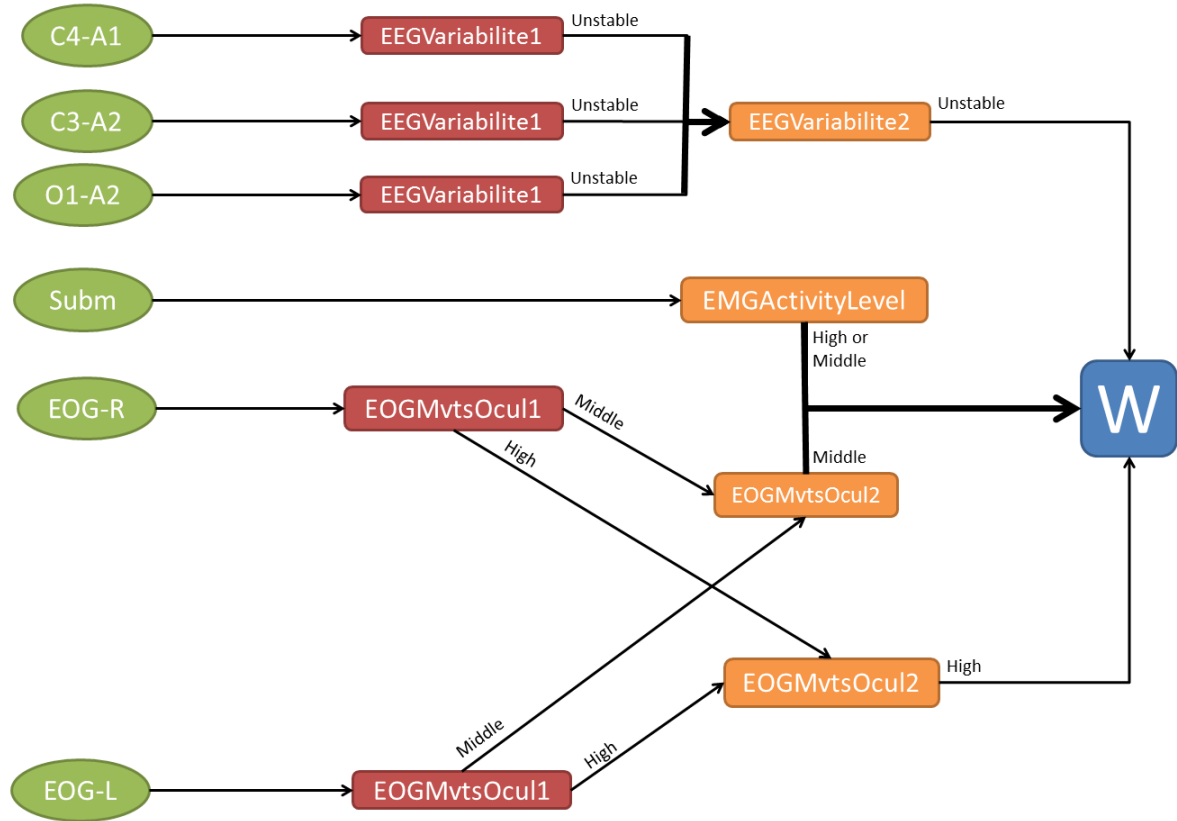


FIGURE 2.34 – Règles de fusion pour tester le stade W



La figure 2.35 illustre les règles pour tester si une époque est du stade REM ou non.

Ainsi, une époque est scorée comme étant de stade REM si les conditions suivantes sont toutes satisfaites :

- $EEGLowWavesEnergy2 = \text{« Low »}$ .
- $EEGSleepSpindles2 = \text{« Absence »}$ .
- $EMGActivityLevel = \text{« Lowest »}$ .
- l'une des deux conditions suivantes est satisfaite :
  - ( $EOG Corr = \text{« Disconjugate »}$  **ET**  $EEG Variabilite2 = \text{« Low »}$ ).
  - $EEGProportionThetaInstFreq2 = \text{« High »}$

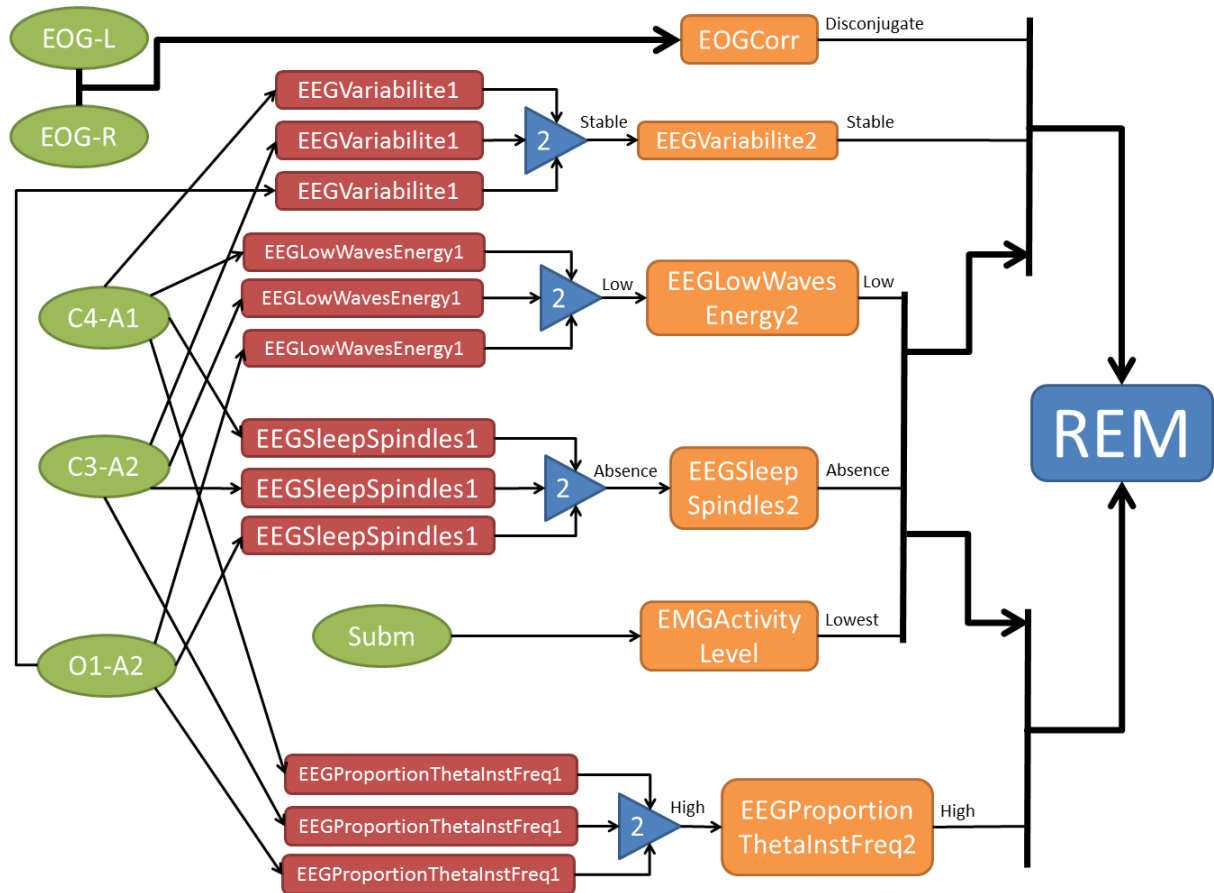


FIGURE 2.35 – Règles de fusion pour tester le stade REM

La figure 2.36 illustre les règles pour tester si une époque est du stade N2 ou non.

Ainsi, une époque est scorée comme étant de stade N2 si l'un des 2 groupes de conditions suivants est satisfait dans sa totalité :

1. – *EEGSleepSpindles2*=« Présence » ;  
 – *EOGMvtsOcul2* IN « Lowest », « Low » ;  
 – *EEGLowWavesEnergy2*=« Low » ;  
 – *EEGProportionThetaInstFreq2*=« High » !.
2. – *EEGSleepSpindles2*=« Présence » ;  
 – *EEGVariabilite2*=« Stable » ;  
 – *EOGCorr*=« Conjugate ».

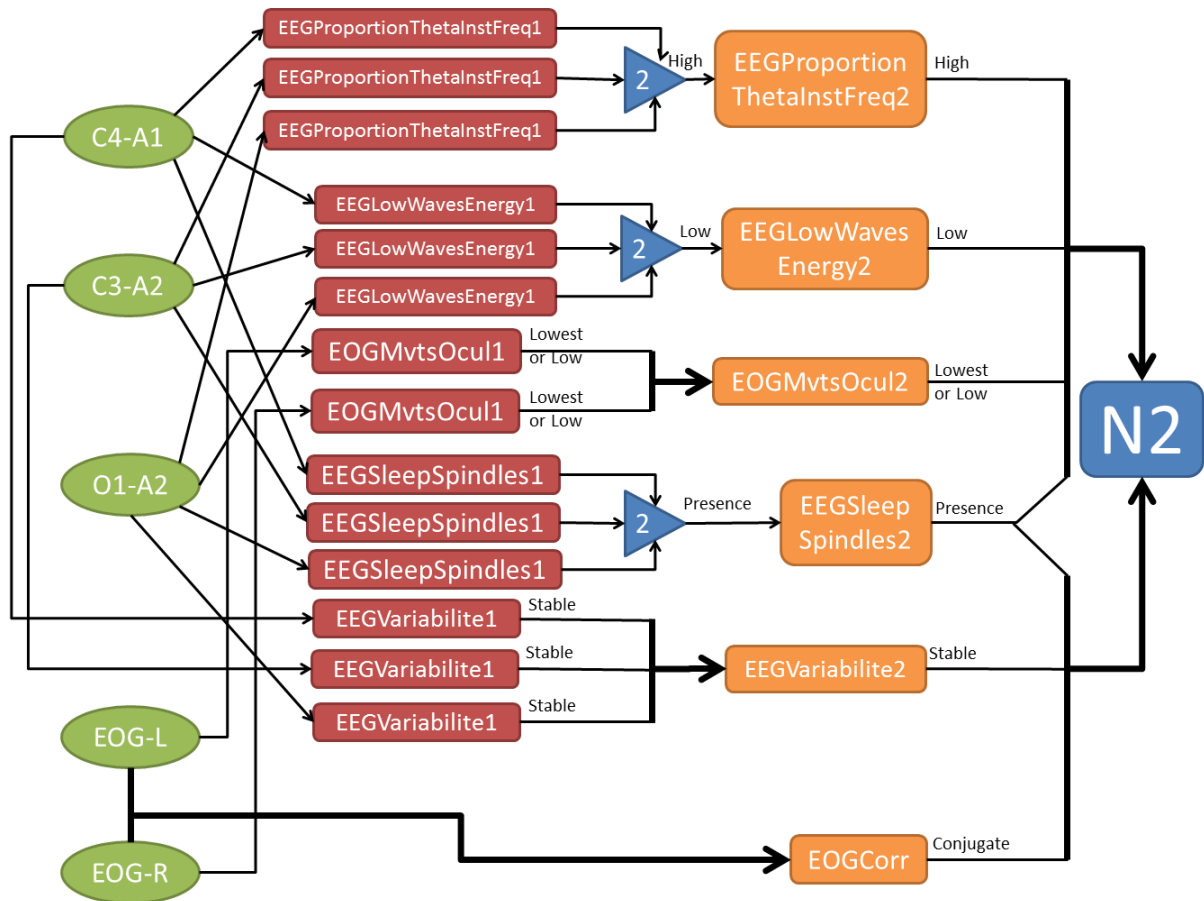


FIGURE 2.36 – Règles de fusion pour tester le stade N2

La figure 2.37 illustre les règles pour tester si une époque est du stade N3 ou non.

Ainsi, une époque est scorée comme étant de stade N3 si l'un des 3 groupes de conditions suivants est satisfait dans sa totalité :

- $EEGVariabilite2 = \text{« Stable »}$  **ET**  $EEGLowWavesEnergy2 = \text{« High »}$  ;
- $EEGSleepSpindles2 = \text{« Absence »}$  **ET**  $EEGProportionLowInstFreq2 = \text{« High »}$  **ET**  $EOGCorr = \text{« Conjugate »}$  ;
- $EEGProportionLowInstFreq2 = \text{« High »}$  **ET**  $EOGMvtsOcul2 = \text{« Lowest »}$ .

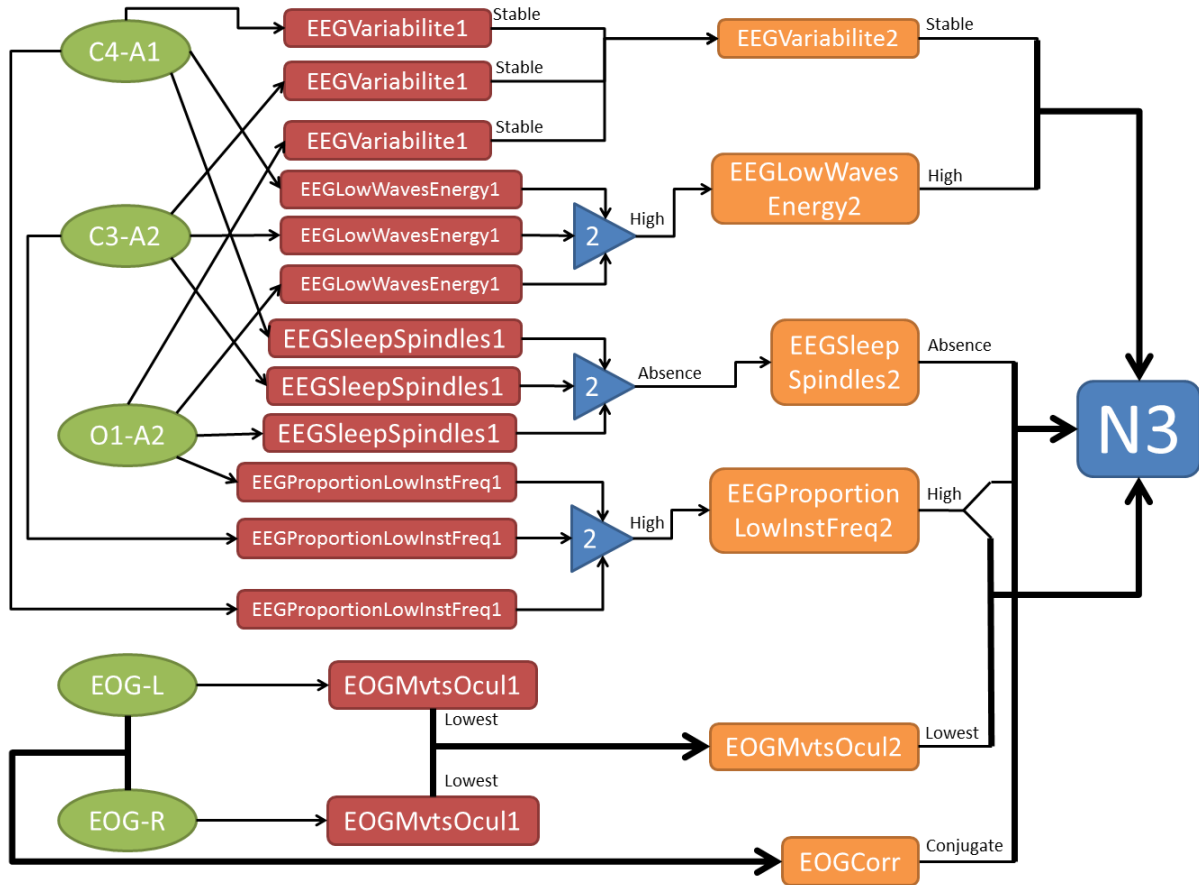


FIGURE 2.37 – Règles de fusion pour tester le stade N3

La figure 2.38 illustre la fusion des paramètres symboliques, époque par époque, de manière à identifier les stades de sommeil et à construire l'hypnogramme.

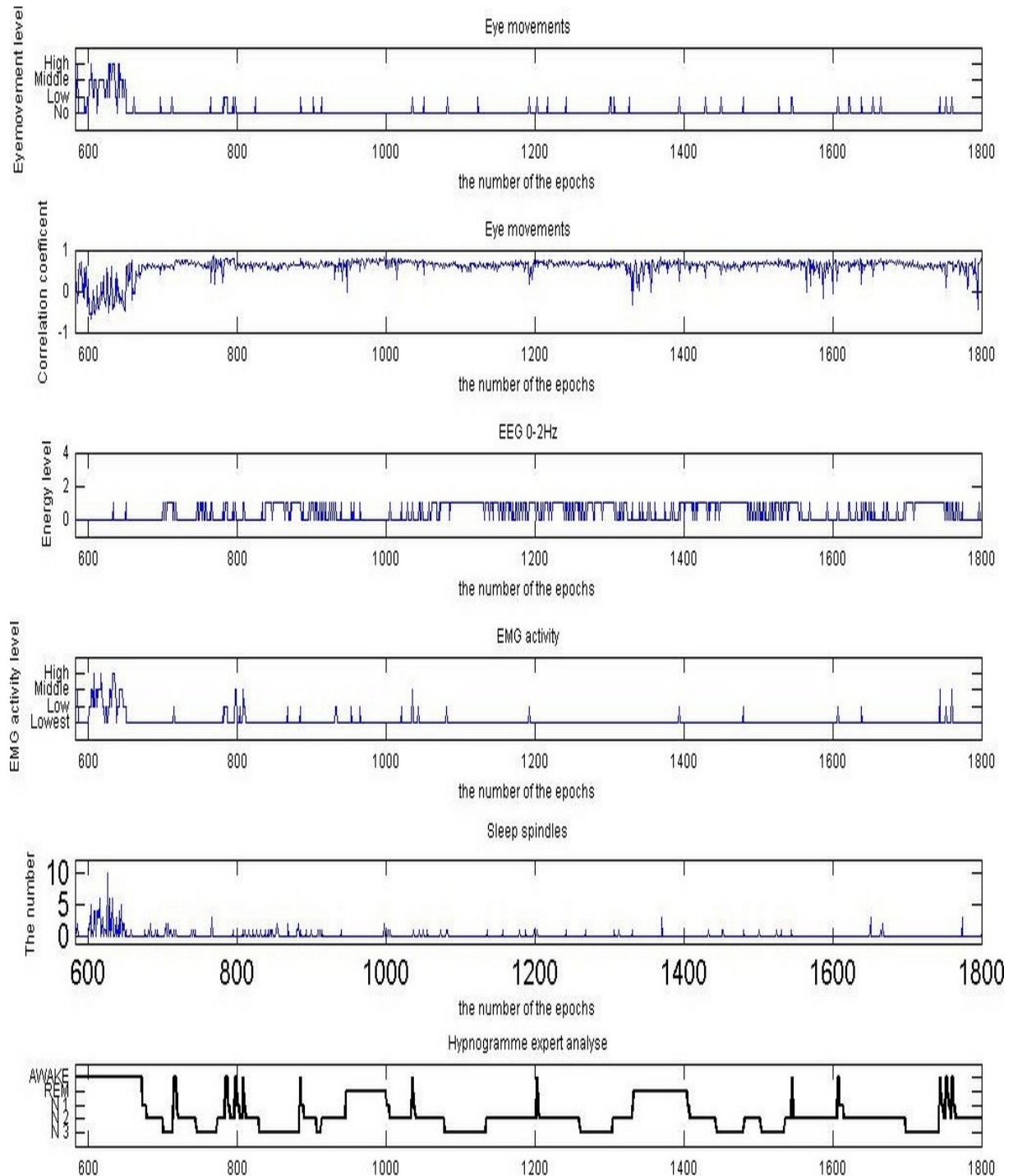


FIGURE 2.38 – Exemple de fusion symbolique

### 3.3 Lissage

Une fois l'hypnogramme obtenu, un lissage est appliqué afin de prendre en compte le voisinage direct de chaque époque et de ne pas avoir de transitions brusques successives et peu vraisemblables entre plusieurs stades de sommeil différents.

Le lissage utilise une fenêtre glissante de 5 époques, centrée sur l'époque  $e$ .

Les règles suivantes sont ensuite appliquées :

- Une époque d'éveil est systématiquement maintenue.
- La valeur affectée à l'époque  $e$  est le stade de sommeil (ou d'éveil) le plus fréquent sur la fenêtre glissante, c'est-à-dire sur un voisinage proche de 2 époques après et avant.

### 3.4 Résultats

Cette méthode a été testée sur 17 patients.

À titre d'exemple, nous proposons sur la figure 2.39 un premier graphique qui compare les hypnogrammes obtenus, avant et après lissage et les compare à l'hypnogramme de l'expert sur un premier patient.

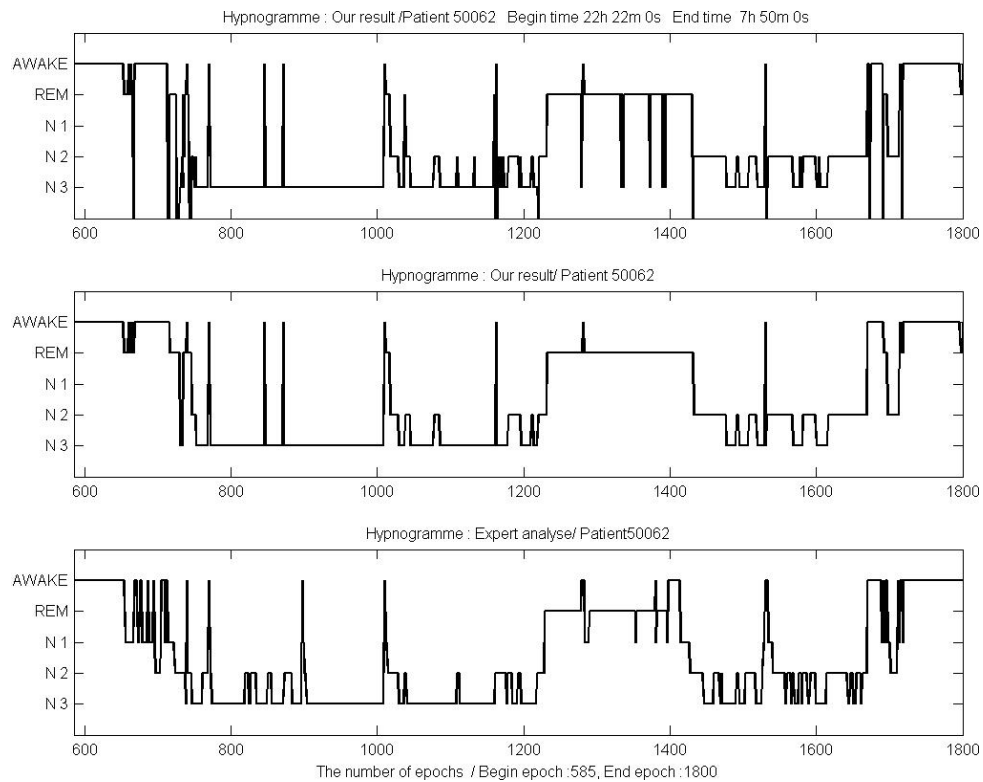


FIGURE 2.39 – Comparaison de l'hypnogramme de l'expert avec ceux obtenus par l'analyse automatique du patient 50062



FIGURE 2.40 – Hypnogrammes obtenus sur les 17 patients tests

La figure 2.40 donne les hypnogrammes obtenus, avant et après lissage mis en vis à vis de l'hypnogramme de l'expert pour les 17 patients. Nous pouvons voir que globalement, les résultats sont assez satisfaisants.

Pour analyser plus en détails ces résultats, observons les taux d'accord obtenus pour chaque stade de sommeil, patient par patient. Ces résultats sont donnés dans le tableau 2.9. Seules les époques incluses dans la période d'analyse de l'expert, c'est-à-dire visualisées et scorées explicitement par le médecin, sont prises en compte dans la comparaison. Ceci représente un total de 21793 époques, soit plus de 181 heures d'enregistrement.

TABLEAU 2.9 – Taux d'accord pour chaque stade de sommeil, patient par patient

ID Patient	Nombre total d'époques	Nombre d'époques analysées	W	R	N1	N2	N3	Total
627	1800	1077	78,35	72,45	0,00	84,86	93,93	83,48
639	1740	1347	100,0	57,53	0,00	4,28	95,95	50,18
1818	1800	1218	58,04	6,35	0,00	41,48	95,98	54,23
3617	1680	1170	86,13	81,74	0,00	84,10	74,65	82,10
3774	1830	1830	69,07	81,44	0,00	97,76	60,87	76,82
3870	1800	1320	83,33	73,73	0,00	62,26	92,45	78,10
3875	1830	1830	83,09	72,09	0,00	84,72	74,19	82,49
3928	1797	1229	81,36	49,21	0,00	82,07	96,30	82,07
4064	1860	1233	100,0	2,67	0,00	89,87	50,58	67,05
4062	1860	1170	15,32	65,28	0,00	66,29	97,39	52,90
45813	1680	1139	88,28	48,72	0,00	27,52	51,06	54,08
46618	1800	1050	100,0	77,95	0,00	51,82	90,28	79,69
48814	1560	1100	41,10	92,00	0,00	53,14	97,19	65,66
50062	1800	1216	84,75	98,04	0,00	57,66	84,14	79,70
50897	1773	1295	74,61	14,58	0,00	24,73	100,0	57,73
55341	1740	1310	63,55	74,38	0,00	49,87	95,87	66,69
56508	1560	1259	58,73	68,09	0,00	85,61	87,63	78,11

Pour faciliter leur interprétation, la figure 2.41 en fait une représentation graphique.

Les résultats sont résumés dans le tableau 2.10.

TABLEAU 2.10 – Intervalle de confiance par stade de sommeil

<b>W</b>	74,45 ± 22,19%
<b>R</b>	60,96 ± 28,48%
<b>N1</b>	0,00 ± 0,00%
<b>N2</b>	61,65 ± 26,50%
<b>N3</b>	84,62 ± 16,44%
<b>Total</b>	70,06 ± 12,18%

Ces résultats sont quasiment comparables à ceux qui seraient obtenus par 2 experts humains interprétant les mêmes données, ce qui est encourageant. En particulier, l'éveil et le sommeil profond (N3) sont bien reconnus. Les taux de reconnaissance du sommeil lent et du sommeil paradoxal sont certes un peu faibles, mais ceci peut s'expliquer par la population étudiée, suspectée de souffrir d'apnée du sommeil dont les troubles surviennent habituellement plus favorablement en sommeil lent et en sommeil paradoxal. Les micro-éveils associés sont des perturbations importantes des signaux neu-

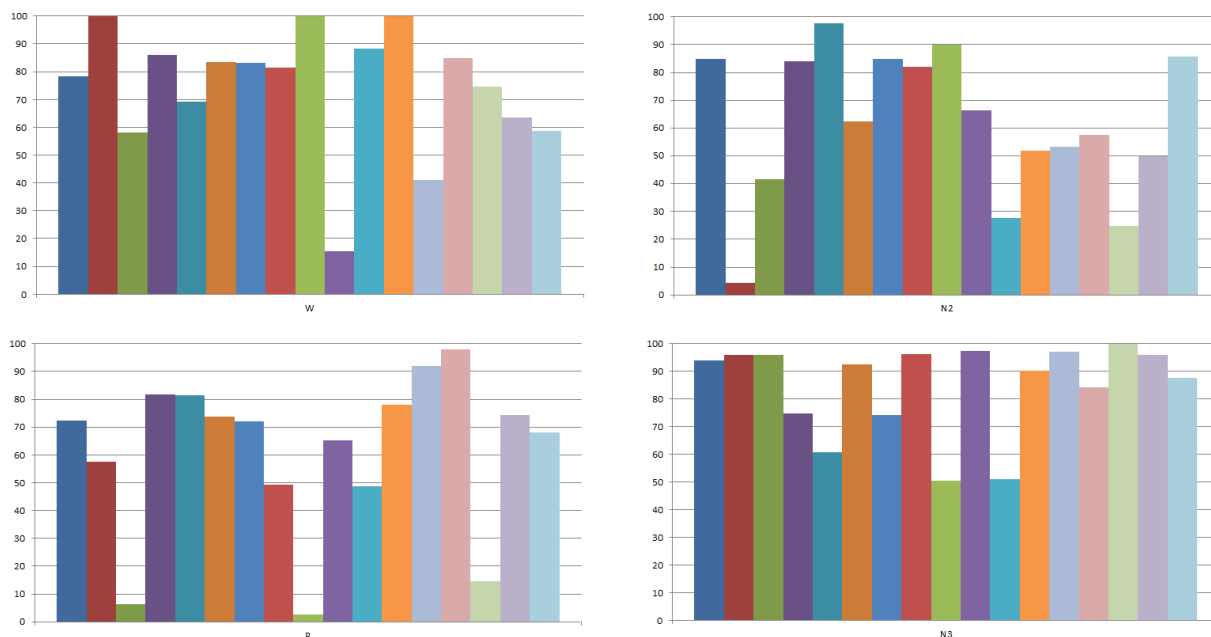


FIGURE 2.41 – Comparaison des résultats pour chaque stade de sommeil

rologiques qui rendent plus difficiles la reconnaissance automatique des stades de sommeil.

À titre de comparaison, nous fournissons dans le tableau 2.11 les intervalles de confiance obtenus par l'algorithme présenté par HANAOKA *et al* dans [30].

 TABLEAU 2.11 – Intervalle de confiance par stade de sommeil de l'algorithme présenté par Masaaki HANAOKA *et al*

W	R	N1	N2	N3
$50,0 \pm 35,2\%$	$75,5 \pm 8,4\%$	$60,2 \pm 3,8\%$	$91,0 \pm 3,3\%$	$50,0 \pm 8,9\%$

Contrairement à notre méthode, cet algorithme est bon pour la reconnaissance du sommeil lent, moyen pour reconnaître le sommeil paradoxal et faible pour identifier le sommeil lent et l'éveil. Ceci peut s'expliquer par la répartition des stades de sommeil au cours d'une nuit. Dans l'ensemble d'apprentissage, le sommeil lent N2 est bien représenté, le sommeil lent et le sommeil paradoxal le sont un peu moins et l'éveil est très peu représenté. Le taux élevé obtenu pour la reconnaissance du stade N1 doit être confronté aux résultats obtenus intra-scoreurs dans [69]. En effet, Michael H. SILBER a montré que lorsque le même expert interprète le même enregistrement à plus de 6 mois d'intervalle, il obtient au mieux 24% d'accord entre les deux interprétations pour le stade N1. Un taux d'accord à 60% serait donc probablement très différent s'il était mesuré sur une interprétation du même enregistrement par le même expert plusieurs mois plus tard. Il n'est pas donc pas très significatif.



### 3.5 Discussion

Pour analyser nos résultats, il s'agit d'étudier les différences de décision entre l'analyse automatique et l'analyse visuelle de l'expert. L'utilisation des paramètres symboliques facilite l'analyse.

La figure 2.42 compare les hypnogrammes de l'analyse automatique et de l'analyse visuelle de l'expert. Il s'en détache une zone mise en évidence que l'expert a identifiée comme sommeil paradoxal R, tandis que l'analyse automatique a conclu à un sommeil lent léger N2. L'analyse des paramètres symboliques permet d'expliquer cette décision. Si le niveau du tonus musculaire est au plus bas, signe caractéristique du sommeil paradoxal, il n'y a pas suffisamment de mouvements oculaires détectés et, en outre, l'analyse a mis en évidence la présence de fuseaux du sommeil, signes aiguillant plus favorablement vers le sommeil lent léger (stade N2) que vers le sommeil paradoxal (stade R). La décision de l'analyse automatique semble donc cohérente.

Chacun des désaccords similaires à celui présenté ci-dessus nécessite une discussion avec l'expert pour comprendre les raisons de sa décision et ajuster l'algorithme, en ajustant les seuils, en créant des règles ou de nouveaux paramètres, ou en ajoutant un prétraitement des courbes de certains patients présentant une caractéristique pouvant perturber l'analyse du signal par des méthodes automatiques et conduire à une mauvaise conclusion.

### 3.6 Conclusion

L'identification des stades de sommeil par construction puis fusion de paramètres symboliques permet d'obtenir des premiers résultats satisfaisants, d'autant plus qu'elle offre la possibilité d'être améliorée par des retours argumentés à l'expert.

Un tel retour à l'expert médical pour chaque époque analysée différemment est une tâche chronophage, qui n'a pas encore été réalisée à ce jour. Il devrait en résulter une amélioration des résultats, déjà encourageants, de cette méthode, qui serait alors inspirée non seulement des règles théoriques diffusées par l'AASM dans [31], mais aussi de l'expérience et la pratique des experts médicaux.

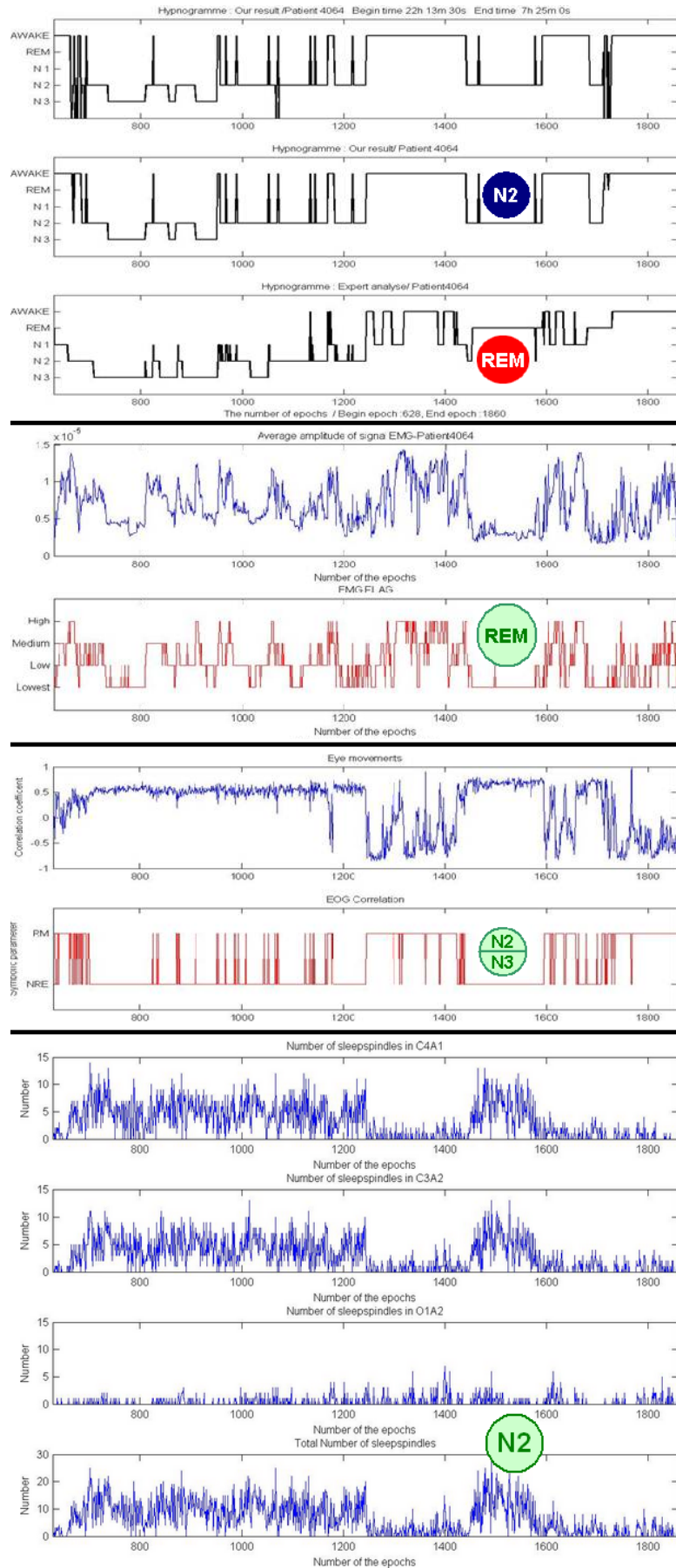


FIGURE 2.42 – Analyse d'un désaccord entre l'expert et l'analyse automatique

## 4 Amélioration d'une interprétation des stades de sommeil par analyse probabiliste

### 4.1 Méthode

Les deux solutions d'interprétation automatique des stades de sommeil proposées précédemment, ont pour objectif de procéder à une analyse, époque par époque, sans prendre en compte l'environnement direct, c'est-à-dire les stades de sommeil attribués aux époques précédant directement l'époque analysée. Il est pourtant bien connu que la succession des stades de sommeil ne se fait pas aléatoirement. Nous proposons ici une méthode, utilisant les hypnogrammes issus de  $N$  analyses automatiques ( $N > 1$ ), les comparant et décidant, pour chaque époque, lequel des  $N$  stades de sommeil attribués est le plus vraisemblable.

Pour cela, nous avons construit, à partir des hypnogrammes des experts médicaux sur 215 patients, une table de probabilité de changement d'un stade de sommeil  $s_1$  vers un stade de sommeil  $s_2$  après  $n$  époques « stables », c'est-à-dire scorées de façon identique. Un extrait de cette table de probabilité est donné sur le tableau 2.12. Il s'agit des probabilités de transition d'une époque de stade  $S_1$  vers une époque de stade  $S_2$  après 10 époques passées en stade  $S_1$ .

TABLEAU 2.12 – Probabilités de transition d'une époque de stade  $S_1$  vers une époque de stade  $S_2$  après 10 époques passées en stade  $S_1$

$\Gamma^*$	N1	N2	N3	R	W
N1	0%	62,5%	2,5%	2,5%	32,5%
N2	30,5%	0%	8,6%	21,9%	39,1%
N3	0%	67,7%	0%	3,2%	29,0%
R	1,9%	44,2%	13,5%	0%	40,4%

Il s'agit ensuite de parcourir les époques une par une et de construire un nouvel hypnogramme à partir des  $N$  hypnogrammes initiaux. Il est considéré qu'au départ, le patient est à l'éveil. Les différents hypnogrammes sont ensuite parcourus jusqu'à ce qu'une époque soit scorée avec un stade de sommeil (N1, N2, N3 ou R) par au moins l'une des analyses. C'est alors ce stade de sommeil qui est choisi (en priorisant N1 à N2, N2 à N3, N3 à R). Notons  $s_1$  ce stade de sommeil et notons  $s_{2a}, s_{2b}, \dots, s_{2k}$  les  $k$  stades de sommeil différents proposés par les  $N$  hypnogrammes pour l'époque suivante. Celle-ci est alors scorée par comparaison des probabilités de transition de  $s_1$  vers  $s_{2a}, s_{2b}, \dots, s_{2k}$  après 1 époque stable en utilisant la table de probabilités de transition présentée ci-dessus. Le stade le plus probable est retenu.

Il est procédé de même pour les époques suivantes en comptant le nombre d'époques stables (c'est-à-dire scorées de façon identique) précédant directement l'époque à décider.

L'hypnogramme final utilise donc l'analyse des caractéristiques des signaux sur l'époque, indépendamment des époques environnantes, mais optimise aussi la vraisemblance de la succession des stades de sommeil.

## 4.2 Résultats

Nous avons testé cet algorithme sur 29 patients. Pour ces 29 patients, nous avons effectué une interprétation automatique avec chacune des 2 méthodes présentées précédemment :

1. Algo 1 : Arbre binaire de décision ;
2. Algo 2 : Fusion de paramètres symboliques (sans l'étape finale de lissage) ;

En moyenne, l'algorithme 1 a scoré correctement 59,05% des époques ; l'algorithme 2 a 53,30% d'accord avec l'expert.

Plus en détails, les taux de reconnaissance par stade de sommeil pour les 2 interprétations initiales, ainsi que pour le résultat de leur fusion par analyse probabiliste, sont donnés dans le tableau 2.13.

TABLEAU 2.13 – Taux d'accord moyen par stade de sommeil des interprétations initiales et de celle issue de leur fusion

Algorithme	W	R	N1	N2	N3
Algo 1	65,45%	67,21%	18,81%	60,00%	84,86%
Algo 2	76,92%	61,84%	1,27%	54,73%	67,91%
Fusion	<b>68,21%</b>	<b>61,18%</b>	<b>12,91%</b>	<b>52,38%</b>	<b>85,70%</b>

Le scorage obtenu par fusion a un taux d'accord moyen de 55,90%.

Il est préférable d'analyser le coefficient  $\kappa$  de Cohen :

- Algo 1 : 0,583 ;
- Algo 2 : 0,491 ;
- **Fusion : 0,536.**

## 4.3 Discussion

Nous pouvons observer, qu'avec 2 scorages initiaux, l'amélioration globale n'est pas nette. Cependant, cela est utile pour améliorer le taux d'accord pour le sommeil profond, le résultat de la fusion étant meilleur que celui des 2 algorithmes initiaux, et permet aussi de rattraper les lacunes de l'Algo 2 sur la reconnaissance du stade N1.

Nous pensons que cette méthode est efficace lorsqu'elle fusionne des interprétations issues de plusieurs algorithmes, notamment si ceux-ci ont un « point fort », c'est-à-dire lorsqu'ils sont chacun performant sur un stade de sommeil particulier.

## 5 Conclusion

Nous avons vu dans ce chapitre deux méthodes de reconnaissance des stades de sommeil qui permettent d'obtenir des résultats quasiment comparables à ceux qui seraient obtenus par un autre expert analysant les mêmes données. La première utilise un arbre binaire de décision élaboré par un apprentissage supervisé ; la seconde utilise

l'extraction puis la fusion de paramètres symboliques. Ces deux algorithmes effectuent un scorage, époque par époque, aveuglement des époques environnantes. Afin de prendre en compte la logique de succession des stades de sommeil d'une époque à celle qui la suit, un algorithme est ensuite appliqué pour améliorer la qualité de l'interprétation ; le calcul de vraisemblance repose sur la large base de données d'interprétations des stades de sommeil sur les enregistrements polysomnographiques de patients différents effectués par un expert.

L'évaluation d'une interprétation automatique souffre du fait qu'il est impossible, à ce jour, de considérer comme objectif fiable à 100% l'interprétation effectuée par un expert médical. En effet, si l'on demandait à un autre expert, ou au même, de procéder à une nouvelle interprétation des mêmes enregistrements, les résultats seraient différents ; la concordance entre les deux experts humains peut même être de seulement 25% pour certains stades de sommeil.



# Chapitre 3

## Reconnaissance d'événements physiologiques impliqués dans le Syndrome d'Apnées du sommeil

### 1 Détection automatique des désaturations

#### 1.1 Description de l'algorithme

Parmi tous les événements symptomatiques du Syndrome d'Apnées du Sommeil, les désaturations sont sans doute les plus simples à détecter. En effet, elles répondent à une définition plus simple : une désaturation est une baisse de 3% au moins sur la courbe de Saturation en Oxygène (exprimée en pourcentage) ; il s'agit donc d'une baisse de 3 points au moins. La difficulté principale vient du fait que la baisse peut ne pas être continue.

Par exemple, la succession de valeurs suivantes de la Saturation devra être interprétée comme une seule désaturation : 98 - 97 - 97 - 96 - 97 - 96 - 95 - 94 - 95 - 97 - 98.

Notre détection des désaturations analyse l'ensemble du signal de la Saturation en Oxygène par fenêtres de 30 secondes. Sur chaque fenêtre sont détectées :

- la valeur maximale de la Saturation en  $O_2$  ;
- la dernière position où ce maximum est atteint (notée  $p_{\max}$ ) ;
- la valeur minimale de la Saturation ;
- la dernière position où ce minimum est atteint (notée  $p_{\min}$ ).

Si l'amplitude entre la valeur minimale et la valeur maximale est au moins égale à 3 et que  $p_{\max}$  est située avant  $p_{\min}$  alors une désaturation est créée, commençant en  $p_{\max}$  et finissant en  $p_{\min}$ .

#### 1.2 Résultats

Afin d'évaluer notre algorithme, nous l'avons testé sur un ensemble de 86 courbes de saturation en Oxygène issues d'enregistrements polysomnographiques toutes interprétées (ou au moins validées) par un expert. Pour chacune de ces courbes, nous avons calculé la précision. Le calcul du *rappel* n'aurait pas été significatif car l'expert ne valide les désaturations que sur les périodes de sommeil, inconnues pour notre analyse.

Sur 11481 désaturations annotées par l'expert, 11198 ont été correctement détectées par notre méthode, soit une précision de 97,54%, ce qui est un très bon taux.

### 1.3 Discussion

Malgré les bons scores obtenus lors des tests, nous avons souhaité comprendre les causes de l'échec à détecter certaines désaturations et avons donc analysé 2 de ces cas parmi les 285 désaturations annotées par l'expert mais non détectées par notre méthode. Les courbes sont fournies sur les figures 3.1 et 3.2.

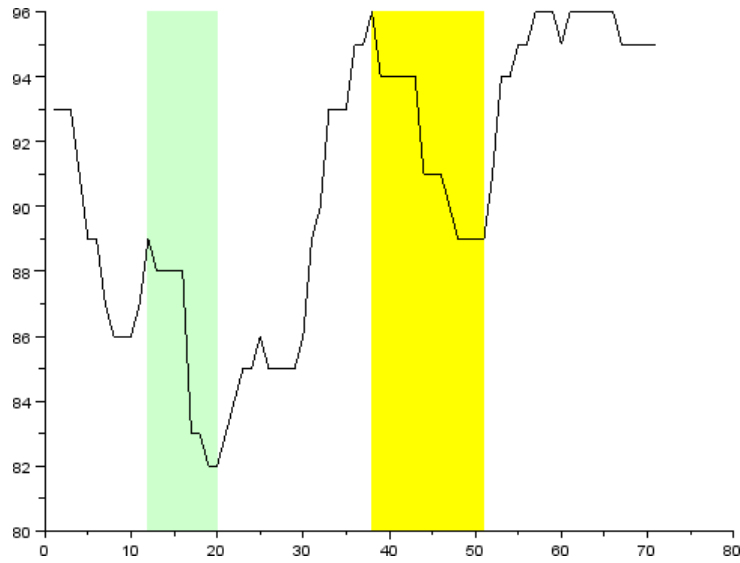


FIGURE 3.1 – Désaturation non détectée n ° 1

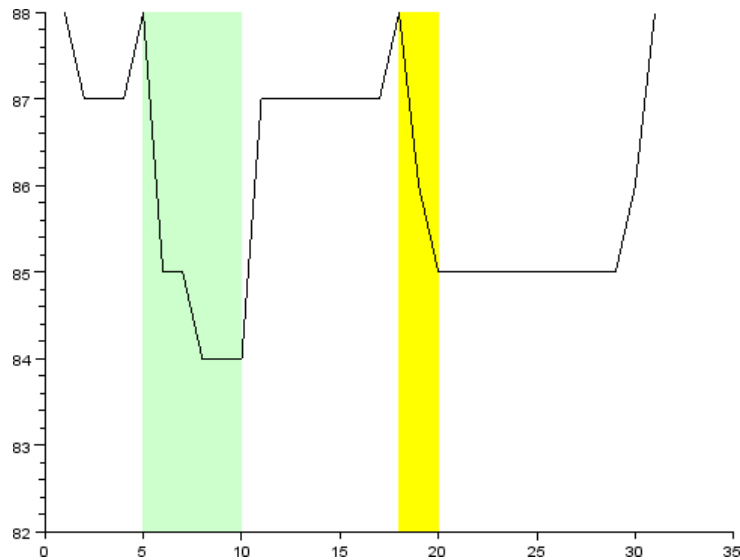


FIGURE 3.2 – Désaturation non détectée n ° 2

La désaturation non détectée est indiquée en jaune. On observe que dans les 2 exemples, elle suit une précédente désaturation plus profonde survenue dans la fenêtre de 30 secondes utilisées pour l'analyse.



Il est possible d'améliorer les résultats en utilisant une fenêtre d'analyse plus courte, mais il est alors possible de rater certaines désaturations assez longues. Par exemple, la désaturation représentée sur la figure 3.3, extraite du même patient que celles des figures 3.1 et 3.2, dure 53 secondes et a une profondeur de 3%. En utilisant une fenêtre trop courte, cette désaturation ne serait pas détectée.

Dans le cadre de cette thèse, le choix a été fait de se « contenter » des résultats présentés ici.

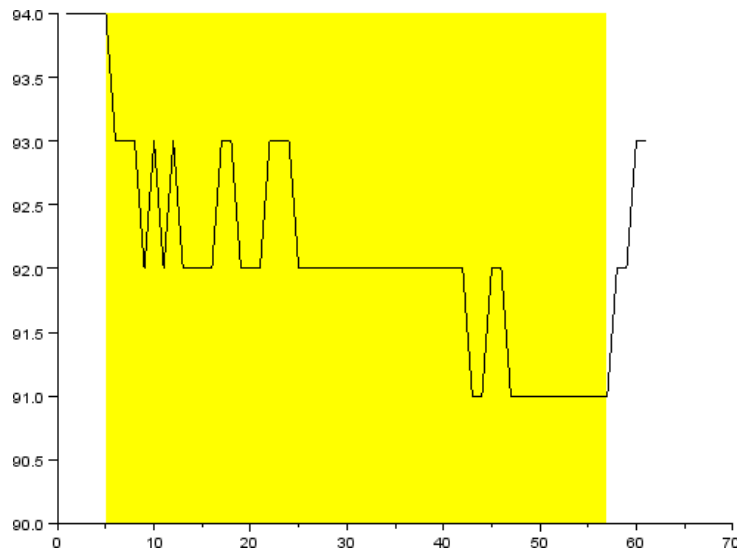


FIGURE 3.3 – Désaturation durant 53 secondes

## 2 Détection automatique des événements respiratoires

La détection automatique des événements respiratoires se déroule en deux étapes successives. Dans un premier temps, on extrait de la courbe représentant le flux nasal les événements ventilatoires triés selon leur nature. Dans un deuxième temps, ils sont classés par type en fonction des résultats sur les courbes représentant l'effort respiratoire.

### 2.1 Reconnaissance des événements ventilatoires selon leur nature

La reconnaissance des événements ventilatoires selon leur nature consiste à extraire de la courbe représentant le flux respiratoire des patterns caractéristiques des apnées ou des hypopnées.

#### Détection des cycles respiratoires

La définition des apnées et des hypopnées se base sur la baisse d'amplitude du flux nasal d'un cycle respiratoire à un autre. Nous avons donc cherché à détecter dans un premier temps l'ensemble des cycles respiratoires afin de les caractériser par différents paramètres.

Ici nous avons considéré comme cycle respiratoire, un cycle constitué d'une partie à pente positive suivie d'une partie à pente assimilée comme nulle suivie d'une partie avec une pente négative.

### 1. Lissage par Moyennes Mobiles ;

La première étape est un lissage de la courbe par moyennes mobiles sur des fenêtres de 31 centièmes de seconde, valeur estimée comme optimale après plusieurs tests.

### 2. Calcul du signe de la pente ;

Il faut au préalable définir une amplitude de référence pour cette étape. Celle-ci est définie comme l'amplitude du signal (flux aérien) sur l'ensemble de l'enregistrement polysomnographique.

La deuxième étape consiste alors à analyser la pente du signal en chaque point de l'enregistrement sur une fenêtre de 3,65 secondes, valeur à nouveau estimée comme optimale après différents tests. Sur la fenêtre, le signal lissé issu de la première étape est analysé selon son amplitude. En dessous du seuil de 0,1‰ de l'amplitude de référence par seconde, c'est-à-dire quand le signal a une amplitude inférieure à 0,365 ‰ de l'amplitude du signal global sur la fenêtre de 3,65 secondes, alors la pente est considérée comme nulle et, par conséquent, le signal constant. Les résultats sont représentés sur la figure 3.4.

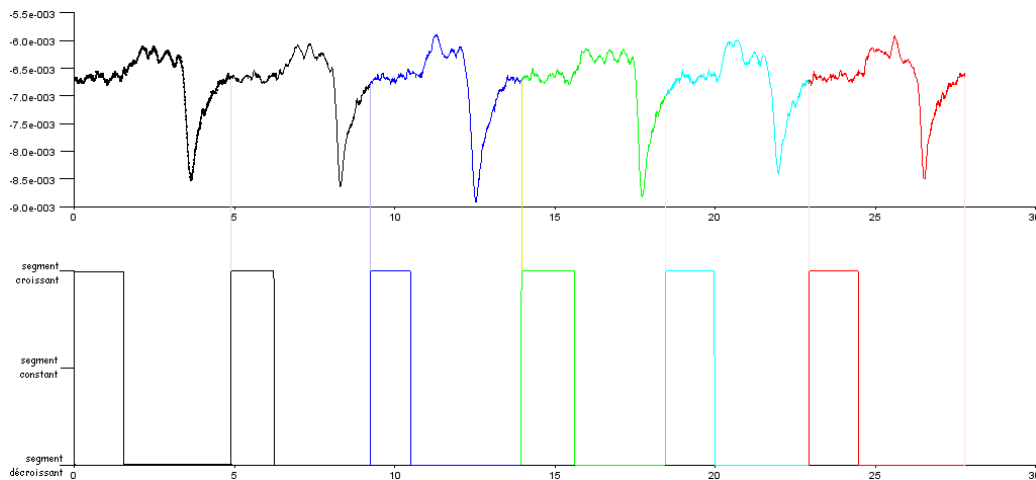


FIGURE 3.4 – Analyse de la pente des Cycles Respiratoires

### 3. Détection des Cycles Respiratoires ;

Les cycles respiratoires sont identifiés comme étant ceux où se succèdent un segment de signal croissant et un signal décroissant, avec éventuellement entre les deux un segment constant ne durant pas plus de 3 secondes.

Cette étape d'identification des cycles respiratoires est difficile à évaluer et valider. Néanmoins, la représentation graphique des résultats montre qu'ils sont bons, comme on peut le voir sur la figure 3.5 où chaque cycle respiratoire identifié sur la courbe de flux respiratoire (c'est-à-dire la pression nasale) est représenté par une couleur différente.

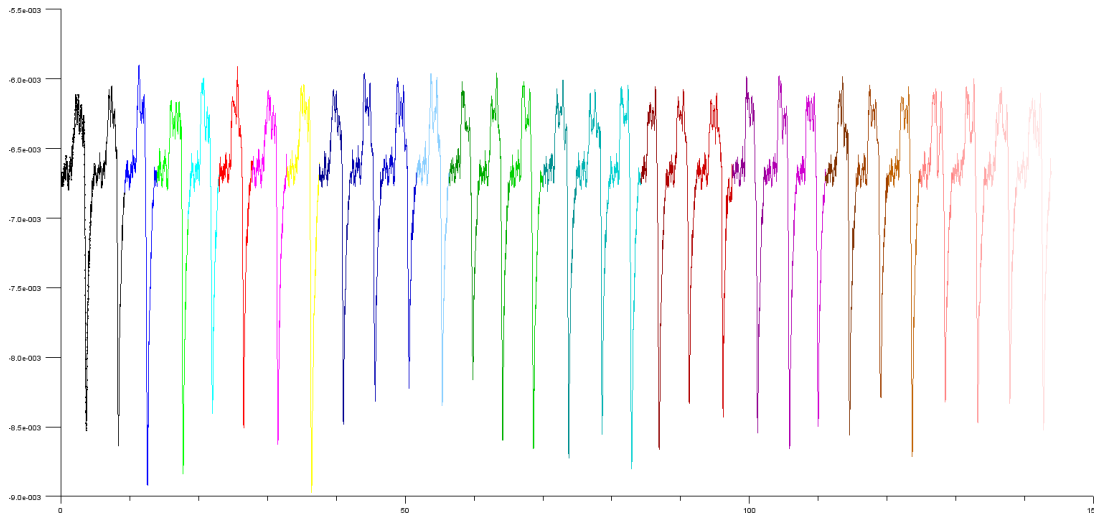


FIGURE 3.5 – Détection automatique des cycles respiratoires

### Classification des événements ventilatoires selon leur nature

Connaissant le début et la fin des cycles respiratoires à partir des résultats de l'étape précédente, un vecteur est construit indiquant l'état respiratoire en chaque point de l'enregistrement.

#### 1. Calcul de l'état respiratoire ;

À partir du quatrième cycle respiratoire détecté sur l'enregistrement, on calcule, pour chaque cycle respiratoire un seuil de référence, valant la moyenne des amplitudes des 3 cycles respiratoires précédents. Si l'amplitude du cycle respiratoire étudié est inférieure à 10 % du seuil de référence, alors l'état respiratoire est noté en « *cessation de flux respiratoire* » ; sinon, si elle est inférieure à 70 % du seuil de référence, alors l'état respiratoire est noté en « *diminution de flux respiratoire* » ; sinon l'état respiratoire est noté comme « respiration normale ».

#### 2. Lissage de l'état respiratoire ;

Ensuite, le vecteur de l'état respiratoire est lissé. À chaque point du vecteur est associé l'état le plus fréquent sur une fenêtre de 10 secondes centrée en ce point.

### Fusion et Filtres Symboliques

La fusion et les filtres symboliques vont permettre d'effectuer des ajustements de précision et, en conséquence l'amélioration, de la détection des événements ventilatoires. Elle utilise les informations associées à 3 concepts, chacun ayant des informations lui étant associées :

1. les cycles respiratoires,
  - ID ;
  - Heure de début ;
  - Heure de fin ;
  - Position en abscisse (heure) du maximum ;
  - Amplitude.

2. les « *diminutions de flux respiratoire* »,
  - ID ;
  - Heure de début ;
  - Heure de fin ;
3. les « *cessations de flux respiratoire* »,
  - ID ;
  - Heure de début ;
  - Heure de fin ;

1. **Calcul du début de chaque événement respiratoire,**

Cette étape consiste en une fusion symbolique. Les débuts des différents événements ventilatoires sont recalculés. Leur nouvelle valeur est issue de la fusion d'informations issues de plusieurs événements : pour calculer la nouvelle valeur d'heure de début pour chaque événement respiratoire, on recherche le dernier cycle respiratoire complet avant le début de l'événement. La position de son maximum indique la nouvelle heure de début de l'événement respiratoire, conformément aux règles de l'AASM [31] et ainsi qu'il est représenté sur la figure 3.6.

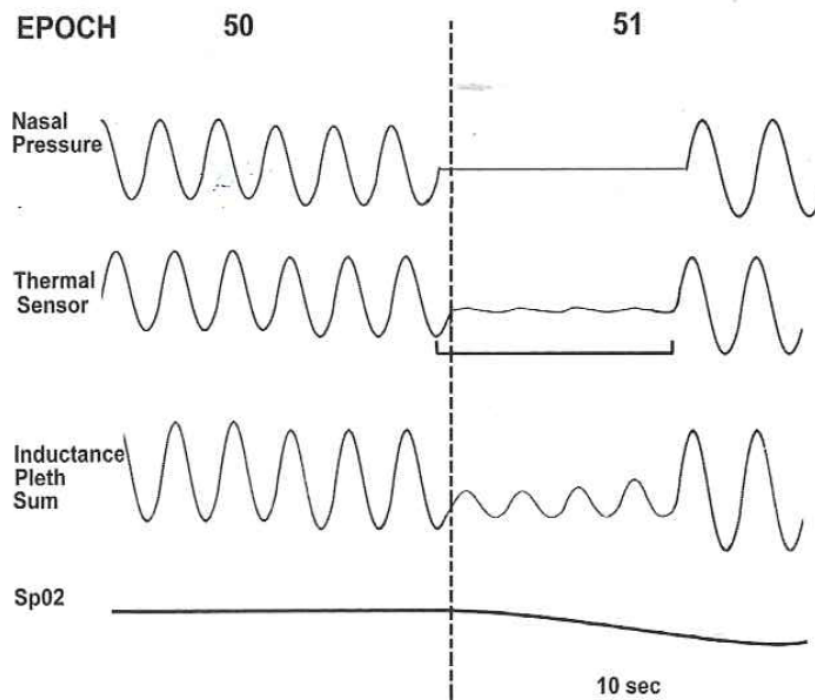


FIGURE 3.6 – Illustration d'une apnée dans la manuel de l'AASM

Prenons par exemple les événements de la figure 3.7 représentés sous la forme de graphes conceptuels. Le graphe conceptuel bleu représente l'événement ventilatoire ; les graphes conceptuels verts représentent 5 cycles respiratoires survenus juste avant ou juste après l'événement ventilatoire. Le dernier cycle respiratoire ayant commencé avant la « *cessation de flux respiratoire* », c'est-à-dire le cycle respiratoire avec l'heure de début la plus tardive parmi tous les cycles respiratoires dont l'heure de début précède l'heure de début de la « *cessation de flux respiratoire* », porte l'ID 2373. La fusion symbolique va permettre de modifier l'heure de début de la

« *cessation de flux respiratoire* » avec pour nouvelle valeur l'heure du max du cycle respiratoire avec l'ID 2373, soit 18/01/2007 - 23h59mn07,55s. Ceci a été représenté sur le graphe conceptuel de la figure 3.8.

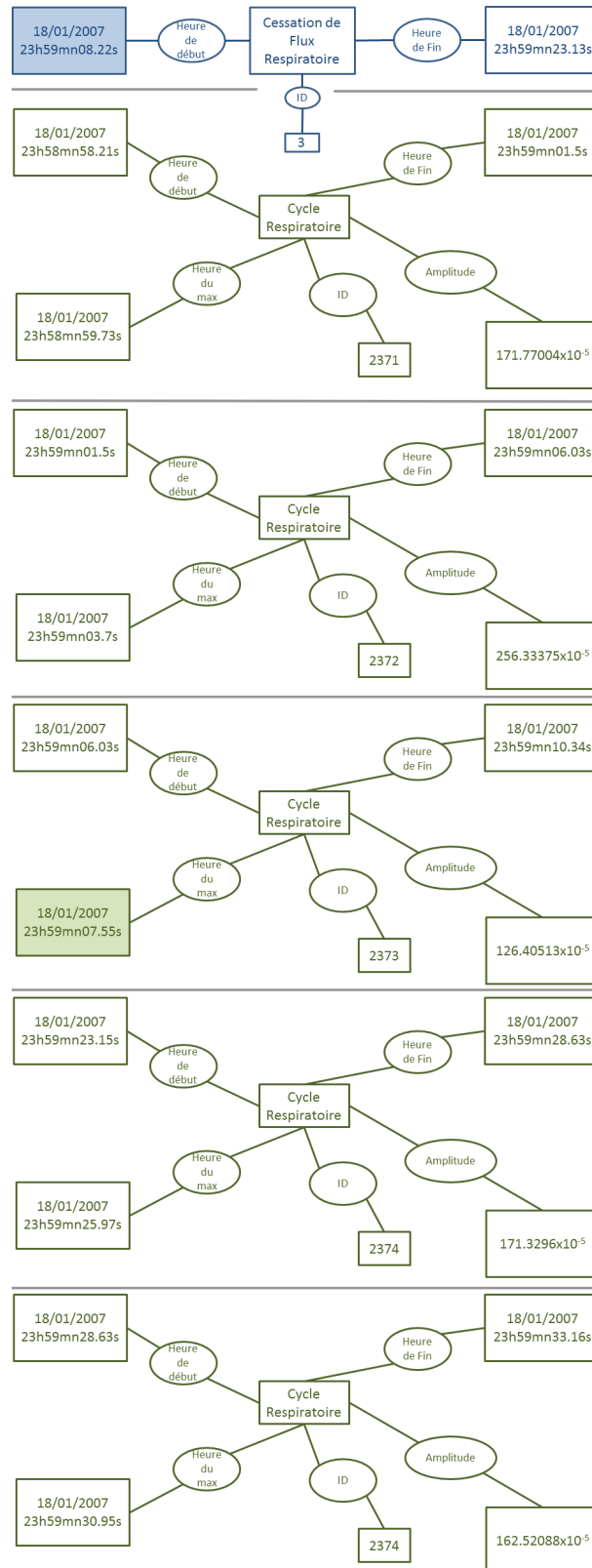


FIGURE 3.7 – Exemple de fusion symbolique entre un événement ventilatoire et un cycle respiratoire

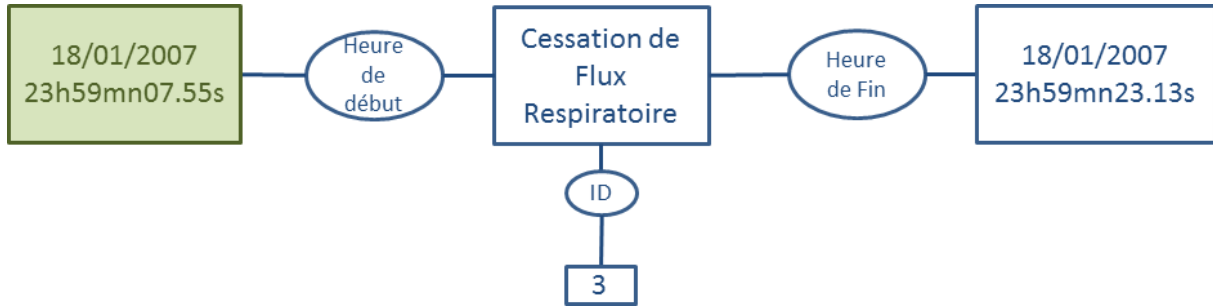


FIGURE 3.8 – Résultat de la fusion symbolique de l'exemple de la figure 3.7

2. **Suppression des événements ventilatoires durant moins de 10 secondes,**  
 Cette étape consiste en un filtre symbolique. Les événements ventilatoires durant moins de 10 secondes sont supprimés. En effet, ils ne peuvent pas répondre aux critères de durée minimale des apnées et des hypopnées définies par l'AASM [31].
3. **Suppression des « diminutions de flux respiratoire » superflues,**  
 Cette dernière étape consiste également en un filtre symbolique, utilisant cette fois ci plusieurs événements. Il y est supprimé des « diminutions de flux respiratoire » superflues, c'est-à-dire ayant répondu à l'ensemble des critères utilisés dans les étapes précédentes, mais n'ayant pas de sens physiologique. C'est le cas où se succèdent, sans pause, une « cessation de flux respiratoire » puis une « diminution de flux respiratoire » ou, dans l'ordre inverse, une « diminution de flux respiratoire » puis une « cessation de flux respiratoire ». La « diminution de flux respiratoire » est alors supprimée et seule la « cessation de flux respiratoire » est conservée.

Par exemple, si on regarde la « cessation de flux respiratoire » (en bleu) et la « diminution de flux respiratoire » (en orange) représentées sous la forme de graphes conceptuels sur la figure 3.9, on constate que la « diminution de flux respiratoire » (ID 197) a lieu à la suite immédiate de la « cessation de flux respiratoire » (ID 4). Elle est donc superflue et doit être supprimée. Seule la « cessation de flux respiratoire » (ID 4) sera conservée parmi ces 2 événements ventilatoires.

À la fin de toutes ces étapes, nous disposons donc de 2 ensembles d'événements ventilatoires :

- les « cessations de flux respiratoire » ;
- les « diminutions de flux respiratoire ».

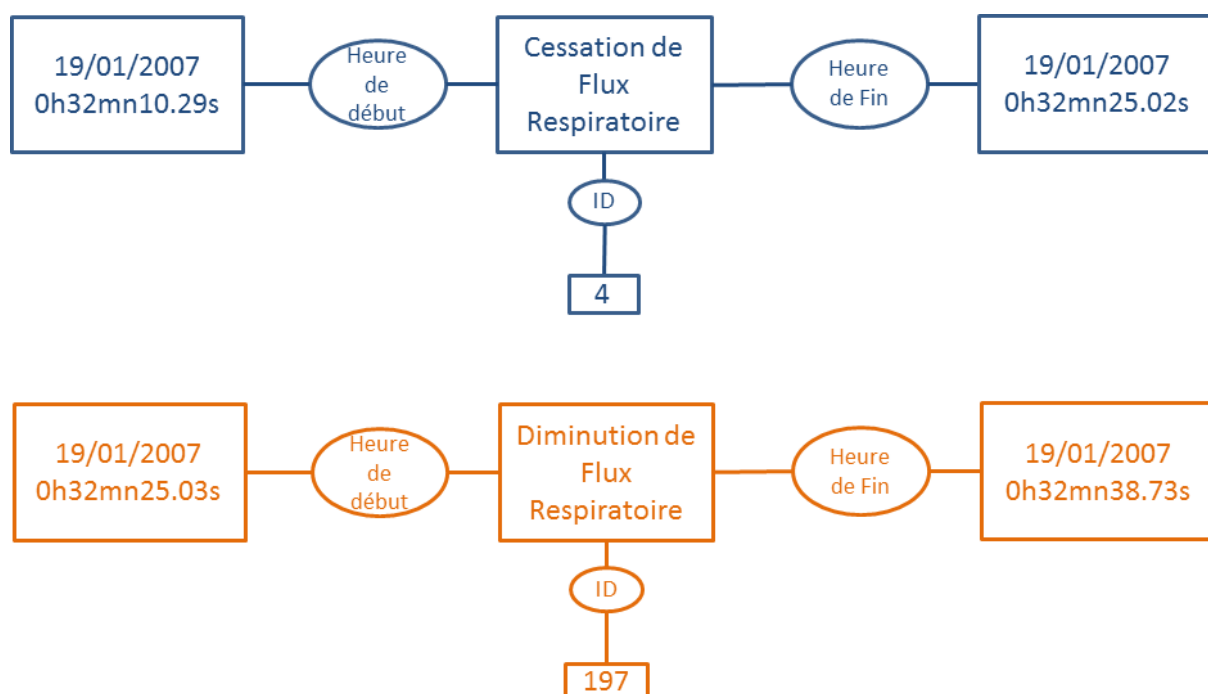


FIGURE 3.9 – Exemple de suppression de « *diminution de flux respiratoire* » superflue

L'algorithme de détection des événements ventilatoires est représenté sur l'organigramme de la figure 3.10.

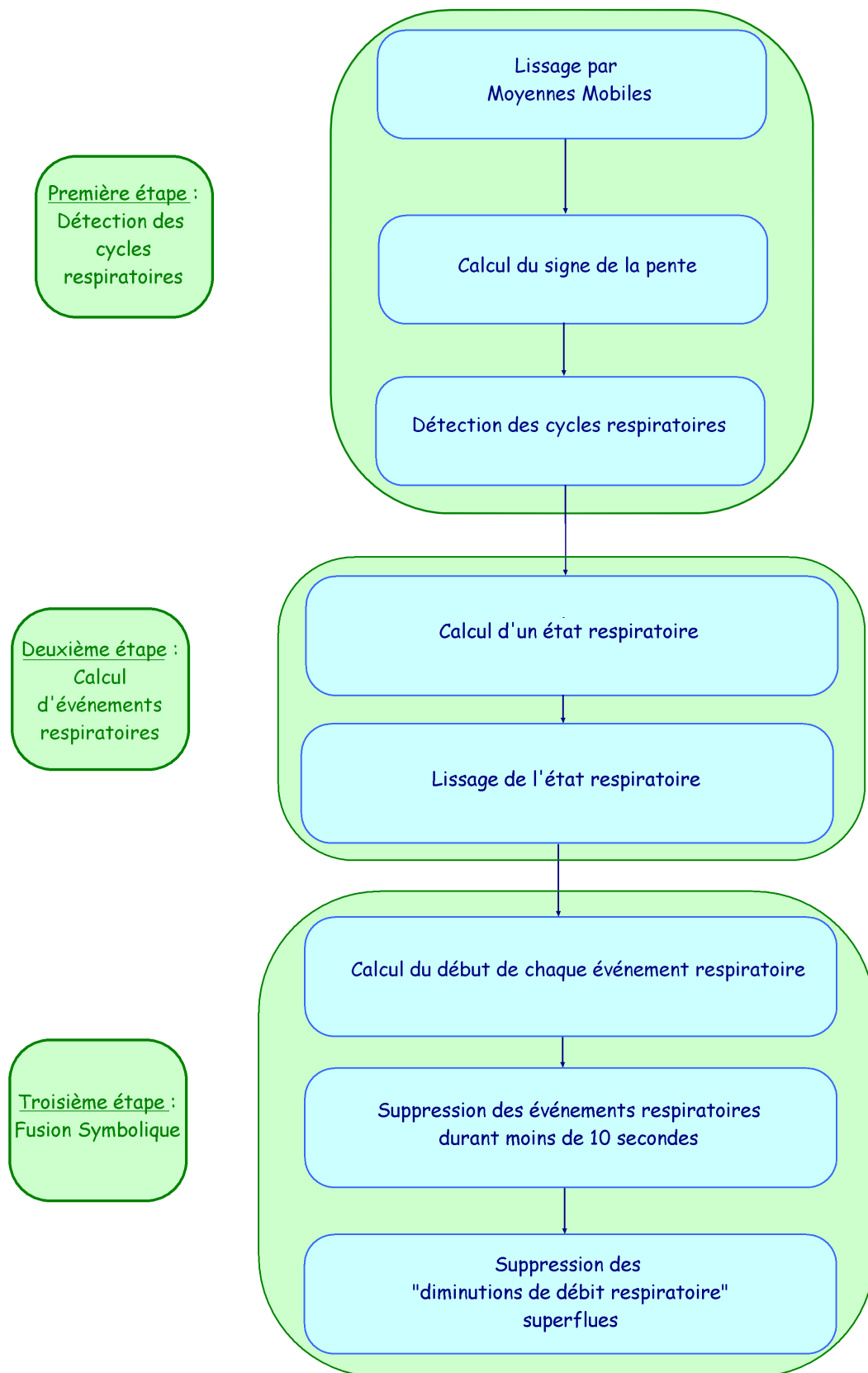


FIGURE 3.10 – Organigramme de l'algorithme de détection des événements ventilatoires



## 2.2 Détection du type d'événement respiratoire par Fusion Symbolique

La deuxième étape consiste à fusionner plusieurs informations symboliques dans le but de répondre aux définitions de l'AASM.

À l'issue de cette étape, nous disposerons des événements suivants :

- apnée obstructive ;
- apnée centrale ;
- apnée mixte ;
- apnée de type indécidable ;
- hypopnée.

À cet effet, nous utiliserons les événements suivants :

- *cessations de flux respiratoire* ;
- *diminutions de flux respiratoire* ;
- désaturations.

En plus, il sera nécessaire d'analyser les courbes de l'effort respiratoire (mesuré par les ceintures thoraco-abdominales) pendant les *cessations de flux respiratoire* pour savoir de quel type est l'apnée.

### Hypopnées

Les règles utilisées pour la détection des hypopnées repose sur la deuxième définition fournie par l'AASM, donnée dans la partie 2.2 du chapitre 1 :

1. Une hypopnée doit durer au moins 10 secondes.
2. Il y a une baisse d'amplitude d'au moins 50% sur le signal de pression nasale pendant au moins 90% de la durée de l'hypopnée.
3. Il y a une désaturation d'au moins 3%.

Les « *diminutions de flux respiratoires* » obtenues à l'étape précédente répondent aux deux premiers points de cette définition. Il reste donc à les fusionner avec le troisième point. Nous allons à nouveau procéder par fusion symbolique.

Pour cela, nous allons suivre la stratégie de fusion suivante : s'il existe, dans l'ensemble des événements disponibles, une « *diminution de flux respiratoire* » et une « *désaturation* » telles que l'heure de début de la « *désaturation* » se situe entre l'heure de début de la « *diminution de flux respiratoire* » et 20 secondes après son heure de fin, ces deux événements sont fusionnés et le résultat de cette fusion est un nouvel événement, une *hypopnée* dont l'heure de début et l'heure de fin sont les mêmes que celles de la « *diminution de flux respiratoire* ».

Un exemple est donné sur la figure 3.11 : la « *diminution de flux respiratoire* » portant l'ID 150, et représentée par le graphe conceptuel bleu portant le numéro 1, peut être fusionnée avec la désaturation portant l'ID 187, et représentée par le graphe conceptuel orange portant le numéro 2, car elle démarre 2 secondes après sa fin, c'est-à-dire dans le

délai de 20 secondes utilisé dans les règles de fusion. Le résultat de cette fusion est un nouvel événement, une hypopnée, représentée par le graphe conceptuel vert portant le numéro 3.

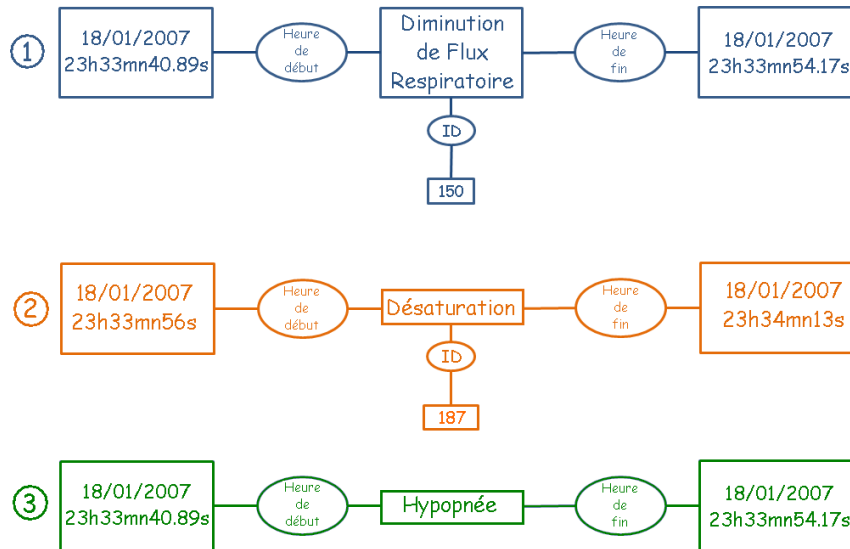


FIGURE 3.11 – Exemple de fusion d'une « diminution de flux respiratoire » et d'une désaturation pour construire une hypopnée

Si, au contraire, il n'existe aucune désaturation commençant dans les 20 secondes après la fin de la « diminution de flux respiratoire », alors la fusion est impossible et aucun nouvel événement n'est créé.

### Détection automatique du type des Apnées

Selon l'AASM, une apnée est définie ainsi (voir partie 2.2 du chapitre 1) :

1. La cessation de flux respiratoire dure au moins 10 secondes.
2. Il y a une baisse d'amplitude du signal d'au moins 90% pendant au moins 90% de la durée de l'apnée.

Les cessations de flux respiratoires obtenues précédemment répondent exactement à cette définition. Il faut ensuite reconnaître le type des apnées détectées, en respectant les définitions de l'AASM (partie 2.2, chapitre 1) :

1. Une apnée est une apnée obstructive si elle est associée à une présence continue ou croissante de l'effort inspiratoire pendant toute la durée d'absence de flux respiratoire.
2. Une apnée est une apnée centrale si elle est associée à une absence d'effort inspiratoire pendant toute la période d'absence de flux respiratoire.
3. Une apnée est une apnée mixte si elle est associée à une absence d'effort inspiratoire dans sa première partie, suivie par une reprise de l'effort inspiratoire dans sa deuxième partie.

Pour effectuer la reconnaissance automatique du type des apnées, nous allons utiliser une méthode de fusion symbolique.

À chaque « *cessation de flux respiratoire* » sont associés deux paramètres liés à la présence d'effort respiratoire. Ces paramètres sont obtenus par une analyse séparée puis comparative des courbes des mouvements thoraco-abdominaux pendant la période de chaque « *cessation de flux respiratoire* ».

Pour pouvoir reconnaître les apnées mixtes, il faut diviser l'apnée en deux parties. Nous avons choisi arbitrairement de travailler sur les deux moitiés de l'apnée ; la coupure se fait alors au milieu de la « *cessation de flux respiratoire* ».

Le premier paramètre est le carré du *coefficient de corrélation linéaire* de Bravais-Pearson [61] lié à la droite des moindres carrés [41], c'est-à-dire la droite de régression linéaire. En effet, en cas d'absence d'effort respiratoire, les muscles respiratoires s'arrêtent et les mouvements thoraco-abdominaux deviennent linéaires, pour revenir à un état d'équilibre physique. Afin de voir si les mouvements sont linéaires, nous allons utiliser la méthode des moindres carrés et construire la droite de régression linéaire. Si on note  $\hat{y} = a \cdot x + b$  l'équation de la droite de régression linéaire (ou droite des moindres carrés) d'un nuage de points, les valeurs de  $a$  (coefficient directeur) et  $b$  (ordonnée à l'origine) sont obtenues grâce aux équations suivantes :

$$\begin{cases} a = \frac{\text{cov}(x,y)}{\text{var}(x)} = \frac{\text{cov}(x,y)}{s^2(x)} \\ b = \bar{y} - a \cdot \bar{x} \end{cases}$$

où :

- $\bar{x} = \sum_{i=1}^n x_i$ , moyenne des  $x_i$  ;
- $\bar{y} = \sum_{i=1}^n y_i$ , moyenne des  $y_i$  ;
- $\text{cov}(x, y) = \frac{1}{n} \times \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})$ , covariance de  $x$  et  $y$  ;
- $s^2(x) = \text{var}(x) = \frac{1}{n} \times \sum_{i=1}^n (x_i - \bar{x})^2$ , variance de  $x$  ;
- $a$  et  $b$  sont les paramètres de la droite de régression ;
- $x_i$  et  $y_i$  sont les points sur la portion de la « *cessation de flux respiratoire* » sur laquelle on applique la régression.

Le coefficient de corrélation linéaire est alors défini ainsi :

$$r_{xy} = \frac{\text{cov}(x, y)}{S_x \times S_y} = \frac{\frac{1}{n} \times \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\frac{1}{n} \times \sum_{i=1}^n (x_i - \bar{x})^2} \times \sqrt{\frac{1}{n} \times \sum_{i=1}^n (y_i - \bar{y})^2}}$$

La paramètre utilisé  $R^2$  vaut donc :

$$R_{xy}^2 = \frac{\left[ \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y}) \right]^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2}$$

où :

- n est le nombre de points du nuage de points sur lequel est effectuée la régression ;
- a et b sont les paramètres de la droite de régression définis ci-dessus ;
- $x_i$  et  $y_i$  sont les points sur la portion de la « *cessation de flux respiratoire* » sur laquelle on applique la régression.

La deuxième mesure utilisée est le coefficient de corrélation entre la courbe des mouvements du thorax et celles des mouvements de l'abdomen. En effet, lorsqu'il y a persistance d'effort respiratoire et que les voies aériennes sont obstruées, les mouvements du thorax et de l'abdomen se désynchronisent. Ceci se manifeste par un déphasage des signaux. Un faible coefficient de corrélation est donc significatif d'une présence d'effort respiratoire. Ce deuxième paramètre est défini par la formule suivante :

$$\text{corr}(Tho, Abd) = \frac{\text{cov}(Tho, Abd)}{S_{Tho} \times S_{Abd}} = \frac{\sum_{i=1}^n (Tho_i - \overline{Tho}) \cdot (Abd_i - \overline{Abd})}{\sqrt{\sum_{i=1}^n (Tho_i - \overline{Tho})^2} \cdot \sqrt{\sum_{i=1}^n (Abd_i - \overline{Abd})^2}}$$

En travaillant sur l'ensemble des apnées relevées par l'expert médical sur un échantillon de patients, nous avons calculé les valeurs de seuils significatifs de la présence et de l'absence d'effort respiratoire pour chacune des 2 courbes (mouvements thoraciques et mouvements abdominaux).

Les seuils utilisés pour chaque paramètre pour considérer qu'il y a **présence** d'effort respiratoire sont les suivants :

- le coefficient de corrélation linéaire des mouvements de l'abdomen sur la première moitié de la « *cessation de flux respiratoire* » est **inférieur à 0,151**.
- le coefficient de corrélation linéaire des mouvements du thorax sur la première moitié de la « *cessation de flux respiratoire* » est **inférieur à 0,086**.
- le coefficient de corrélation entre les courbes des mouvements de l'abdomen et du thorax est **inférieur à -0,212**.

Les seuils utilisés pour chaque paramètre pour considérer qu'il y a **absence** d'effort respiratoire sont les suivants :

- le coefficient de corrélation linéaire des mouvements de l'abdomen sur la première moitié de la « *cessation de flux respiratoire* » est **supérieur à 0,514**.
- le coefficient de corrélation linéaire des mouvements du thorax sur la première moitié de la « *cessation de flux respiratoire* » est **supérieur à 0,499**.
- le coefficient de corrélation entre les courbes des mouvements de l'abdomen et du thorax est **supérieur à 0,854**.

En utilisant ces prédicats, nous pouvons raisonner et décider s'il y a, globalement, présence ou absence d'effort respiratoire, ou s'il est impossible de le déterminer. Nous considérons qu'il y a présence (resp. absence) d'effort respiratoire sur l'une des moitiés de la « *cessation de flux respiratoire* » si, sur cette portion, le coefficient de corrélation linéaire atteste de présence (resp. absence) d'effort respiratoire pour chacune des courbes **ou** si le coefficient de corrélation des mouvements de l'abdomen et du thorax atteste de présence (resp. absence) d'effort respiratoire. Entre les seuils, il est considéré que c'est indécidable.

Ensuite, les règles suivantes sont appliquées pour décider du type de chaque apnée :

- S'il y a mesure de présence d'effort respiratoire sur les deux moitiés de la « *cessation de flux respiratoire* », alors on considère que l'apnée est une apnée **obstructive**.
- S'il y a mesure d'absence d'effort respiratoire sur les deux moitiés de la « *cessation de flux respiratoire* », alors on considère que l'apnée est une apnée **centrale**.
- S'il y a mesure d'absence d'effort respiratoire sur la première moitié, mais présence sur la deuxième moitié alors l'apnée est une apnée **mixte**.
- Dans les autres cas, l'apnée sera considérée comme **indécidable**.

## 2.3 Résultats

Nous avons testé notre algorithme de détection des événements respiratoires et de reconnaissance de leur type sur un ensemble de 86 enregistrements polysomnographiques tous intégralement interprétés par un expert médical.

### Détection des événements ventilatoires

Dans un premier temps, nous avons testé la détection des événements respiratoires sans leur type, c'est-à-dire la détection des « *cessations de flux respiratoire* » et des « *diminutions de flux respiratoire* ».

Sur l'ensemble de ces données, les experts médicaux ont relevé un total de 3748 apnées et 7362 hypopnées, c'est-à-dire un total de 11110 événements respiratoires. Parmi eux, certains ont été détectés et correctement identifiés, c'est-à-dire que les apnées ont été identifiées comme des « *cessations de flux respiratoire* » et les hypopnées comme des « *diminutions de flux respiratoire* ». D'autres événements ont été également correctement reconnus, mais mal identifiés, c'est-à-dire que les apnées ont été identifiées comme des « *diminutions de flux respiratoire* » et les hypopnées comme des « *cessations de flux respiratoire* ». Enfin, certains événements n'ont pas du tout été détectés.

Les résultats que nous avons obtenus sont présentés dans le tableau 3.1.

Globalement, nous avons donc correctement détecté et identifié 59,82% des apnées et 68,4% des hypopnées relevées par l'expert médical et correctement détecté mais mal identifié 28,5% des apnées et 10,9% des hypopnées. Enfin, 11,7% des apnées et

TABLEAU 3.1 – Classification des apnées et des hypopnées relevées par l'expert médical

	événement	Classification par l'analyse automatique		
		<i>cessations de flux respiratoire</i>	<i>diminutions de flux respiratoire</i>	non détectés
Expert	Apnées	2242	1068	438
	Hypopnées	806	5035	1521

20,7% des hypopnées n'ont pas été détectées par notre méthode. Ceci est illustré sur le graphique de la figure 3.12. Il est à noter que 94 apnées (soit 2,5%) et 183 hypopnées (soit 2,5%) relevées par l'expert recouvraient plusieurs événements ventilatoires détectés par l'analyse automatique. En cas de recouvrement, les événements ventilatoires de même type prévalaient sur les événements ventilatoires de type différent. Ceci signifie que les apnées ont été comparées prioritairement aux « *cessations de flux respiratoire* » puis aux « *diminutions de flux respiratoire* », et que les hypopnées ont été comparées prioritairement aux « *diminutions de flux respiratoire* » puis aux « *cessations de flux respiratoire* ».

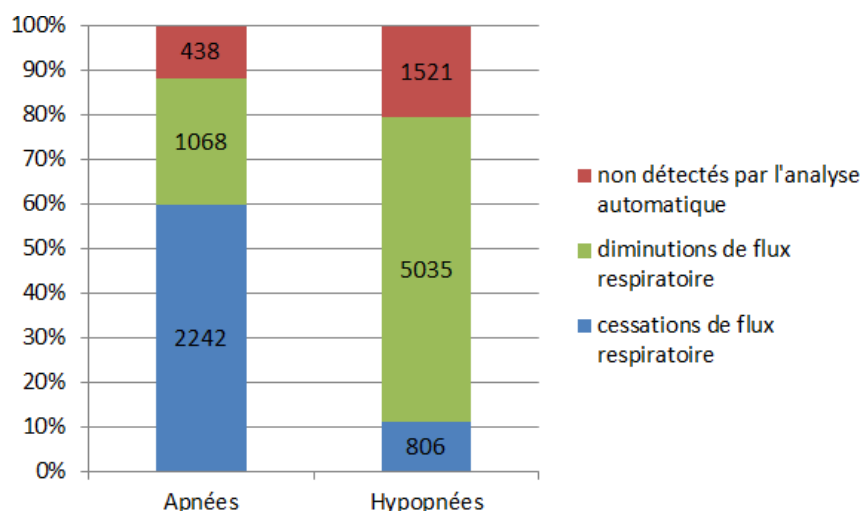


FIGURE 3.12 – Répartition par classe des événements respiratoires relevés par l'expert médical

Si on se place désormais du point de vue de l'analyse automatique, au total, 6103 « *cessations de flux respiratoire* » et 18247 « *diminutions de flux respiratoire* », soit un total de 24350 événements ventilatoires, ont été détectés. Parmi elles, 1088 « *cessations de flux respiratoire* » et 5519 « *diminutions de flux respiratoire* » avaient lieu en dehors de la période d'analyse des données de l'expert. Il reste donc 5015 « *cessations de flux respiratoire* » et 12728 « *diminutions de flux respiratoire* » survenant pendant la période d'analyse de l'expert, soit un total de 17743 événements ventilatoires. À nouveau, parmi ces événements ventilatoires, certains ont été relevés et identifiés de façon identique par l'expert, c'est-à-dire que les « *cessations de flux respiratoire* » avaient été identifiées comme des apnées et les « *diminutions de flux respiratoire* » comme des hypopnées. D'autres événements ont été également relevés par l'expert, mais identifiés différemment, c'est-à-dire que les « *diminutions de flux respiratoire* » avaient été identifiées comme des apnées et les « *cessations de flux respiratoire* » comme des hypopnées. Enfin, certains événements n'avaient pas du tout été détectés par l'expert.

Les résultats sont présentés dans le tableau 3.2.

TABLEAU 3.2 – Classification selon l'expert des « cessations de flux respiratoire » et des « diminutions de flux respiratoire » détectées par l'analyse automatique

	événement	Classification de l'expert			
		Apnées	Hypopnées	non relevés	hors période
Analyse automatique	cessations de flux respiratoire	2243	750	2022	1088
	diminutions de flux respiratoire	751	4113	7864	5519

Globalement, 74,9% des « cessations de flux respiratoire » détectées par l'analyse automatique et relevées comme événement respiratoire, c'est-à-dire scorées en apnées ou hypopnées, par l'expert médical, sont effectivement des apnées, donc des « cessations de flux respiratoire », selon l'expert. De la même façon, 84,6% des « diminutions de flux respiratoire » relevées par l'expert comme événement respiratoire, sont effectivement des hypopnées, donc des « diminutions de flux respiratoire », selon l'expert médical.

Par contre, 25,1% des « cessations de flux respiratoire » avaient également été détectées par l'expert médical, mais étiquetées comme des hypopnées. Pareillement, 15,4% des « diminutions de flux respiratoire » avaient aussi été détectées par l'expert médical, mais étiquetées comme des apnées.

Enfin, sur la période d'analyse de l'expert, 40,3% des « cessations de flux respiratoire » et 61,8% des « diminutions de flux respiratoire » n'avaient pas du tout été relevées par l'expert médical.

Ces résultats sont illustrés sur le graphique de la figure 3.13.



FIGURE 3.13 – Répartition par classe des événements respiratoires relevés par l'analyse automatique

### Reconnaissance du type des événements respiratoires

Nous avons évalué notre méthode de reconnaissance automatique du type des apnées sur le même ensemble de données, c'est-à-dire sur les enregistrements polysomnographiques de 86 patients suspectés de souffrir d'un Syndrome d'Apnées du Sommeil, tous interprétés par un expert médical.

Afin d'illustrer les résultats, nous proposons :

- en figure 3.14 et 3.15 deux apnées obstructives détectées par notre méthode ;
- en figure 3.16 et 3.17 deux apnées centrales détectées par notre méthode ;
- en figures 3.18 et 3.19 deux apnées mixtes détectées par notre méthode.

Sur chacune de ces six figures, sont représentées 3 courbes :

1. La première est la courbe du flux respiratoire. On y voit en rouge la portion du signal correspondant à la « *cessation de flux respiratoire* » telle qu'elle a été détectée par notre analyse automatique.
2. La deuxième et la troisième courbe représentent les mouvements thoraciques et abdominaux. Les deux couleurs utilisées permettent de repérer le milieu de la « *cessation de flux respiratoire* », césure utilisée dans notre analyse automatique. Les droites de moindres carrés sont également représentées afin d'illustrer les différences observées sur notre premier paramètre utilisé (coefficient de corrélation linéaire).

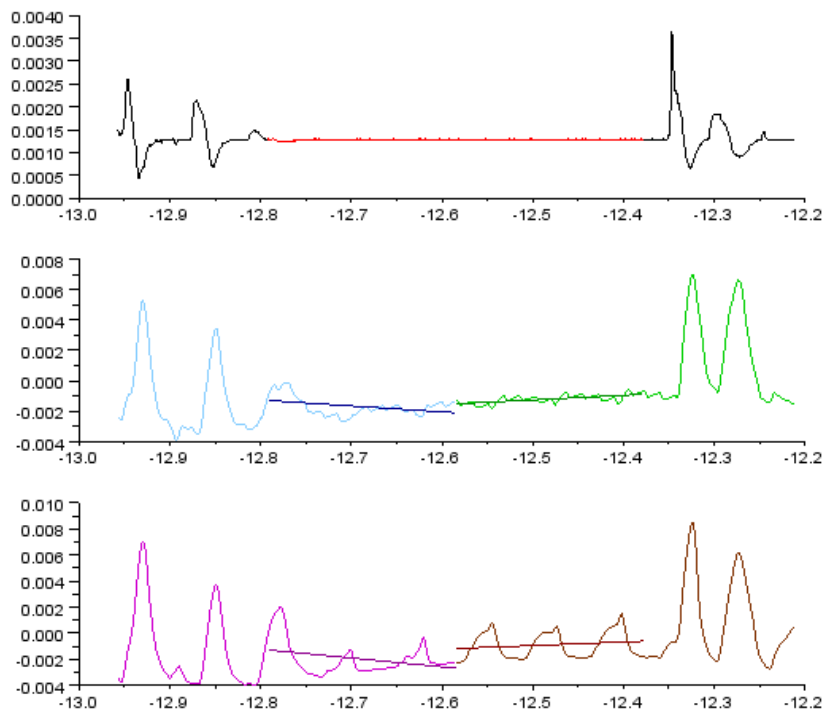


FIGURE 3.14 – Apnée obstructive n ° 1



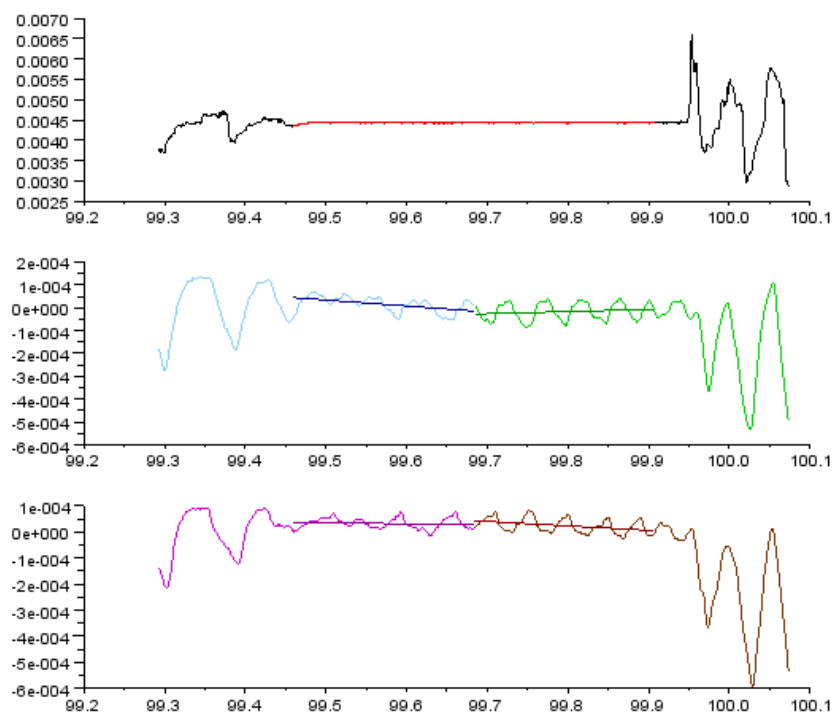


FIGURE 3.15 – Apnée obstructive n ° 2

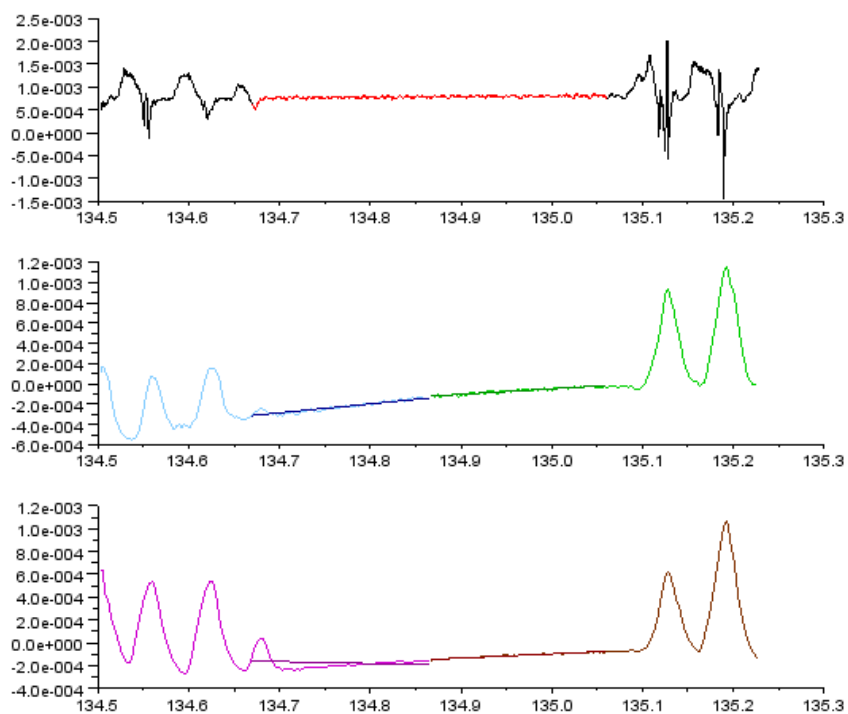


FIGURE 3.16 – Apnée centrale n ° 1

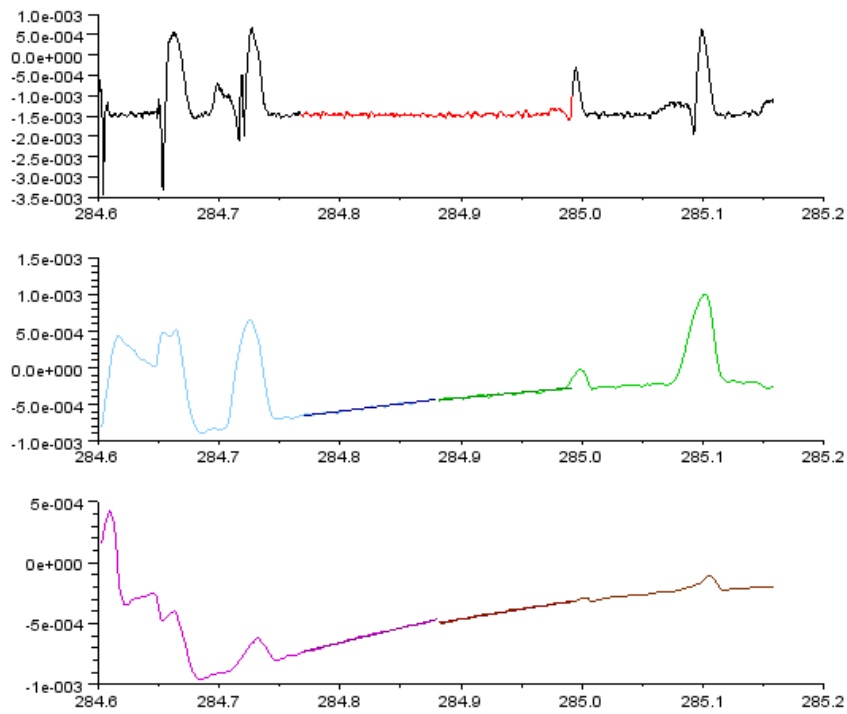


FIGURE 3.17 – Apnée centrale n ° 2

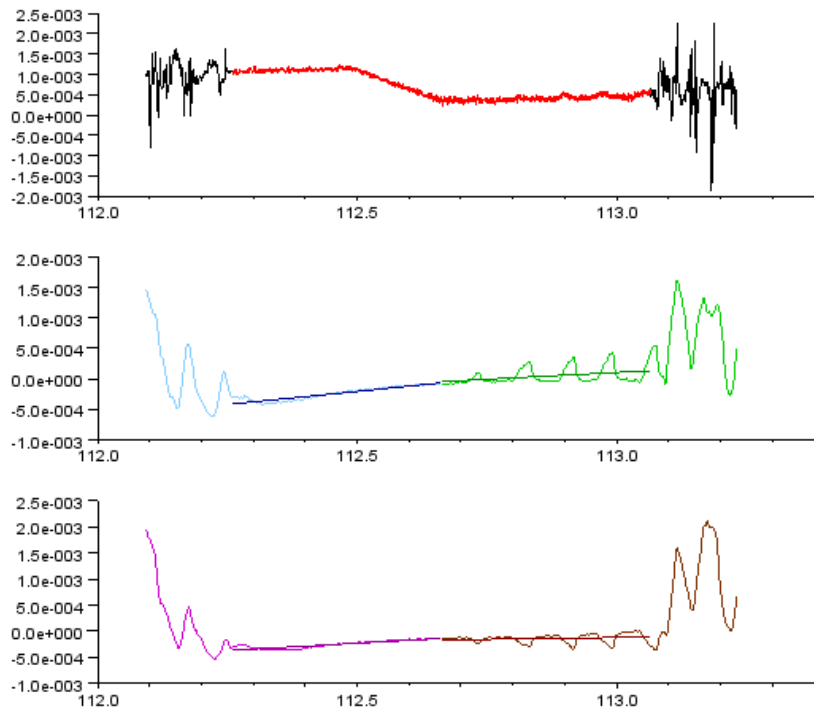


FIGURE 3.18 – Apnée mixte n ° 1

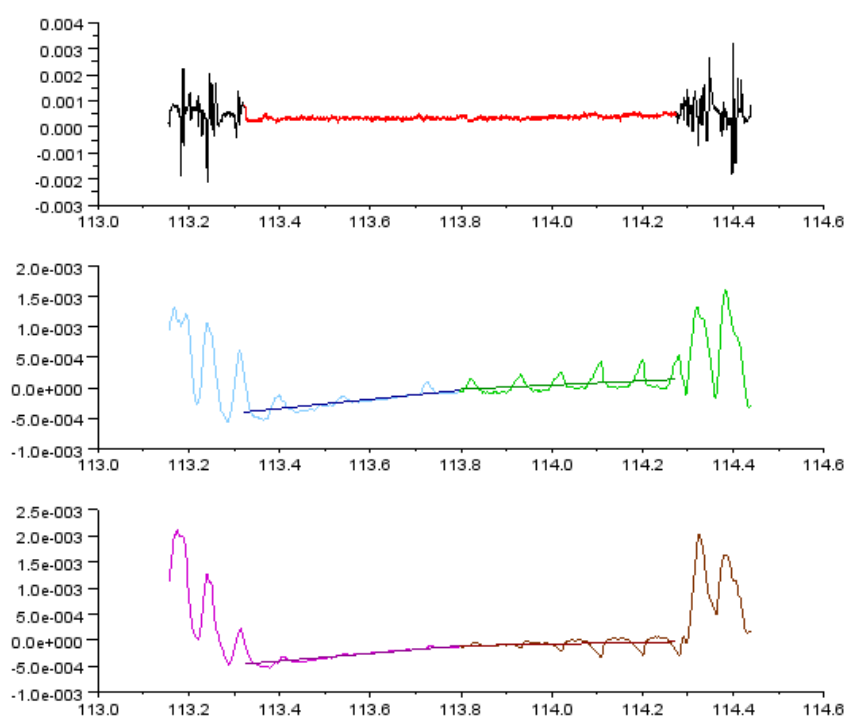


FIGURE 3.19 – Apnée mixte n ° 2

Nous avons les premières statistiques suivantes :

- 52% d'apnées obstructives ;
- 22% d'apnées centrales ;
- 4% d'apnées mixtes ;
- 22% d'apnées de type indécidable.

Ces résultats sont illustrés sur la figure 3.20.

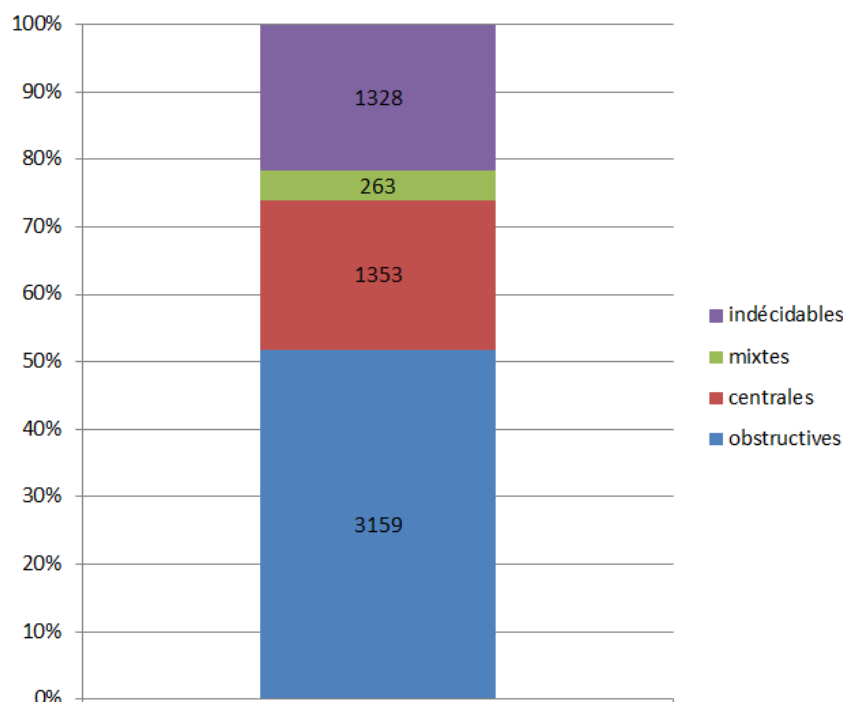


FIGURE 3.20 – Reconnaissance du type des apnées

Nous avons ensuite comparé nos résultats à ceux de l'expert médical.

Dans un premier temps, analysons le rappel pour chaque type d'apnées. Selon l'expert, il y avait 3158 apnées obstructives, 488 apnées centrales et 108 apnées mixtes sur l'ensemble de l'enregistrement, c'est-à-dire un total de 3754 apnées. La petite différence entre le décompte des apnées avec et sans distinction de type s'explique par quelques recouvrements déjà évoqués.

Dans le tableau 3.3, nous pouvons voir le résultat de l'analyse de chacune de ces apnées avec une répartition par type. Nous avons vu que 2,5% des apnées et 2,5% des hypopnées relevées par l'expert médical recouvraient plusieurs événements détectés par l'analyse automatique. En cas de recouvrement, le type obstructif prévalait sur le type mixte, qui prévalait lui-même sur le type central, puis sur le type indécidable et enfin sur les hypopnées.

Ces résultats sont illustrés sur la figure 3.21.

TABLEAU 3.3 – Comparaison de la reconnaissance automatique du type des apnées avec la décision de l'expert

	Types d'apnées	analyse automatique					
		apnées				hypopnées	non détectées
		obstructives	centrales	mixtes	indécidables		
Analyse Expert	obstructives	1443	92	83	392	838	310
	centrales	32	95	8	49	193	111
	mixtes	21	9	8	11	40	19

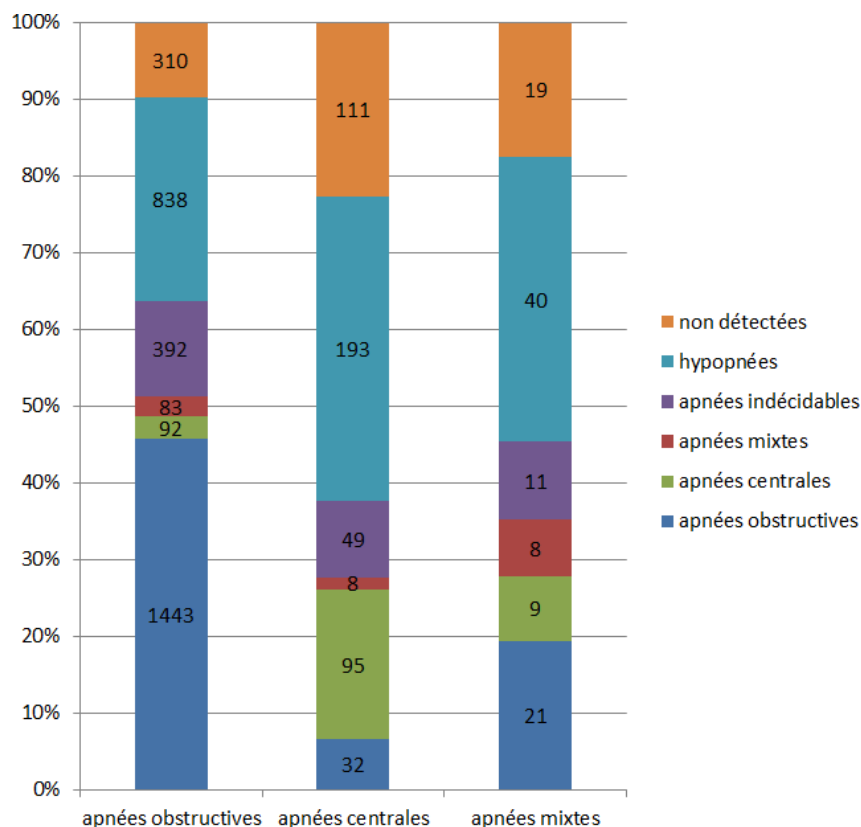


FIGURE 3.21 – Comparaison des résultats de la reconnaissance du type des apnées avec l'interprétation de l'expert médical - Rappel

Mesurons maintenant la précision. Nous avons mesuré que parmi les 5015 apnées détectées par l'analyse automatique, 74, soit 1,5%, recouvrent plusieurs événements respiratoires scorés par l'expert médical. De même, 354 hypopnées, soit 2,8%, parmi les 12728 détectées par l'analyse automatique recouvrent plusieurs événements respiratoires relevés par le médecin. En cas de recouvrement, le type obstructif prévalait sur le type mixte, qui prévalait lui-même sur le type central. Les résultats sont présentés dans le tableau 3.4 et illustrés sur la figure 3.22.

En se restreignant aux apnées détectées et étiquetées comme telles à la fois par l'analyse automatique et par l'expert médical, nous obtenons les résultats suivants :

- 96,6% d'accord avec l'expert pour les apnées étiquetées comme obstructives ;
- 48,7% d'accord avec l'expert pour les apnées étiquetées comme centrales ;
- 8,1% d'accord avec l'expert pour les apnées étiquetées comme mixtes.

TABLEAU 3.4 – Comparaison de la reconnaissance automatique du type des apnées avec la décision de l'expert médical - Précision

	Types d'apnées	Analyse de l'expert				
		apnées			hypopnées	non relevées
		obstructives	centrales	mixtes		
Analyse automatique	obstructives	1438	30	21	442	823
	centrales	91	95	9	113	720
	mixtes	83	8	8	41	86
	indécidables	400	11	49	154	393

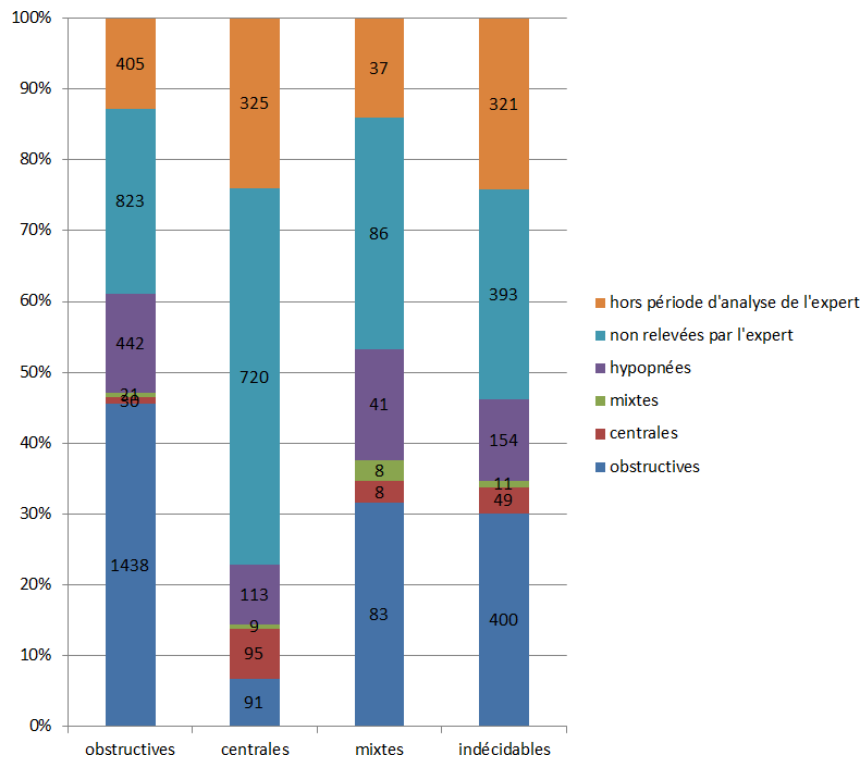


FIGURE 3.22 – Comparaison des résultats de la reconnaissance du type des apnées avec l'interprétation de l'expert médical

## 2.4 Discussion

### Détection des événements ventilatoires

Notre méthode permet une détection assez précise des événements ventilatoires puisque nous avons correctement détecté près de 90% des apnées et plus de 80% des hypopnées. Parmi les événements non détectés, il y a un ensemble d'événements qui ne respectent pas *stricto sensu* les caractéristiques définies par l'AASM, en particulier certains événements respiratoires durent légèrement moins de 10 secondes, ce qui exclut une détection par l'analyse automatique respectant les règles de l'AASM, bien qu'ils puissent avoir un sens physiologique justifiant les choix de l'expert. À partir du fichier d'export de l'interprétation visuelle de l'expert, nous avons mesuré que c'est le cas pour 42 apnées (1,1%) et 47 hypopnées (0,6%).

Le défaut principal de notre méthode se situe sur le fait qu'un nombre très important des événements ventilatoires détectés par l'analyse automatique n'avaient pas été jugés

pertinents par l'expert. Ceci peut s'expliquer par plusieurs raisons. La raison principale est que nous avons travaillé sur l'ensemble de l'enregistrement du flux respiratoire, sans nous restreindre, comme l'expert, à la période de sommeil, différente, a priori, de la période d'analyse de l'expert. Ceci sera corrigé lors de l'étape suivante qui va fusionner les événements obtenus lors de cette étape entre eux, et en particulier, avec les stades de sommeil. Une deuxième raison tient à notre interprétation des règles de l'AASM. En effet, les apnées et les hypopnées sont définies, entre autres, par des caractéristiques sur l'amplitude du signal qui est comparé à un seuil de référence calculé en fonction de l'amplitude des derniers cycles respiratoires. Cette définition a ses limites. En effet, après une apnée, le patient peut être en hyperventilation avec des cycles respiratoires de grande amplitude. Si on utilise cette amplitude comme seuil de référence, l'analyse automatique peut alors générer, de façon erronée, une « *diminution de flux respiratoire* » voire une « *cessation de flux respiratoire* ». De plus, cette définition sous-entend que les cycles respiratoires de référence doivent avoir eu lieu juste avant l'apnée ou l'hypopnée, ce qui n'est pas le cas, dans notre algorithme, après une apnée. Nous avons donc parfois pris en compte, pour le calcul du seuil de référence, l'amplitude de cycles respiratoires survenus une minute avant l'apnée, ce qui a conduit à des erreurs.

Afin de mieux comprendre nos erreurs et d'améliorer notre algorithme, nous proposons en figure 3.23 une sélection d'apnées et d'hypopnées que notre algorithme n'a pas détecté. Chacune des vignettes est constituée de trois courbes : le flux respiratoire, en haut, et, en dessous, les mouvements thoraco-abdominaux.

Nous pouvons y remarquer qu'une partie des événements respiratoires n'a pas été détectée à cause du bruit particulièrement important présent sur le signal du flux respiratoire. Plusieurs événements sont aussi difficiles à identifier sans l'aide des courbes des mouvements thoraco-abdominaux, ce qui peut motiver à étendre notre méthode en utilisant ces 2 courbes dès les premières étapes de reconnaissance des événements ventilatoires.

Cependant, dans le cadre de cette thèse, nous avons décidé de nous contenter d'une précision assez faible, en imaginant un système de support à la décision où l'expert aurait ensuite à valider ou non les éléments détectés lors d'une préanalyse. Toutefois cette analyse permet d'envisager des solutions pour améliorer la qualité de la détection automatique des événements respiratoires.

Il convient également de s'interroger sur la pertinence des apnées détectées par l'analyse automatique mais non relevées par l'expert médical. Pour cela, nous allons utiliser une sélection de ces apnées présentée en figure 3.24. La plupart des événements détectés semblent pertinents.

Enfin, il convient d'analyser les événements qui ont été étiquetés différemment par l'analyse automatique et par l'expert. En regardant, sur la figure 3.25, une sélection d'hypopnées ayant été analysées comme des apnées par l'analyse automatique, on constate qu'il y a effectivement des hypopnées sans ambiguïté, et qu'une grande partie semble répondre aux critères de définition des apnées selon l'AASM.

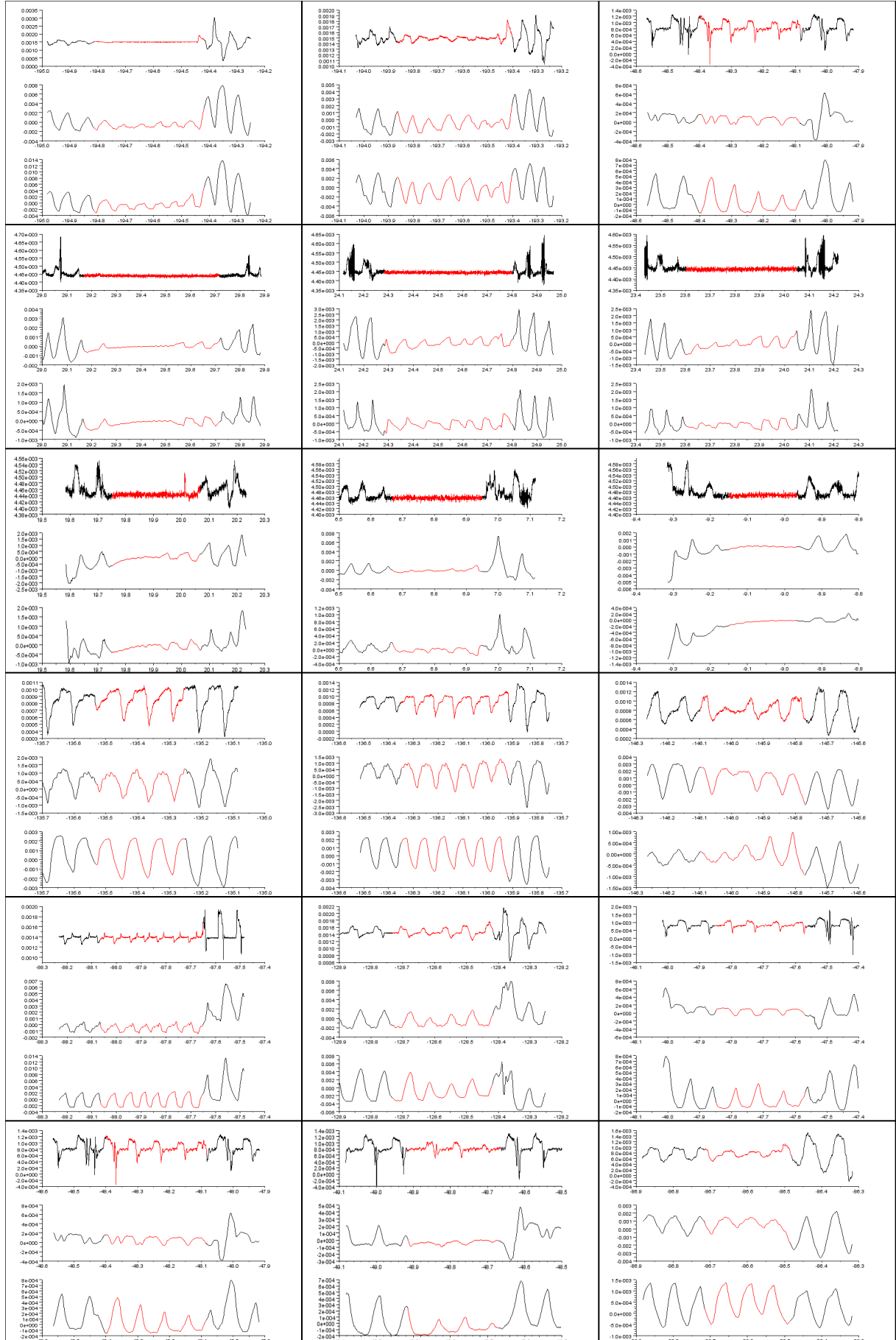


FIGURE 3.23 – Sélection d'apnées et d'hypopnées non détectées par l'analyse automatique



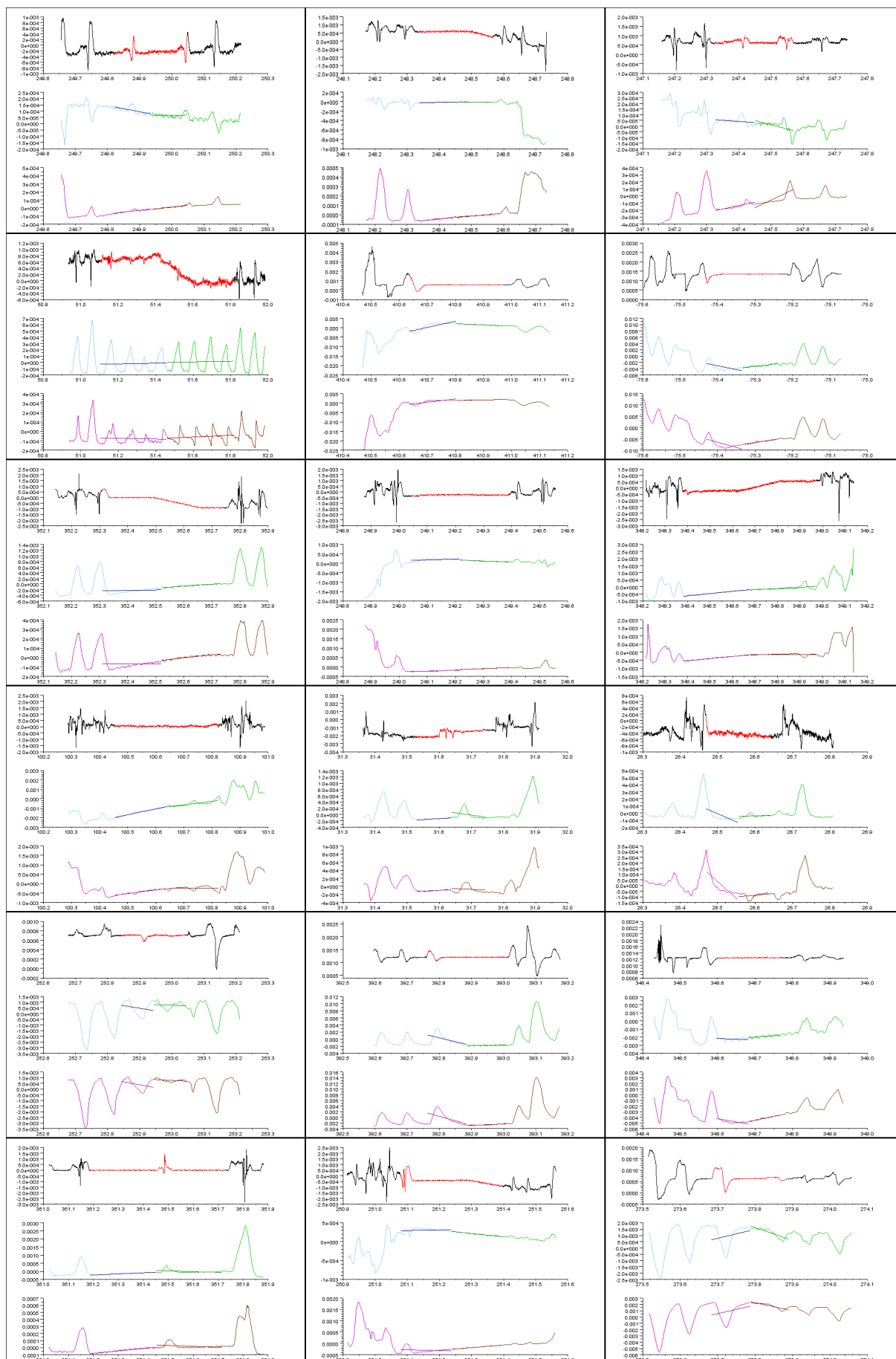


FIGURE 3.24 – Sélection d'apnées non relevées par l'expert

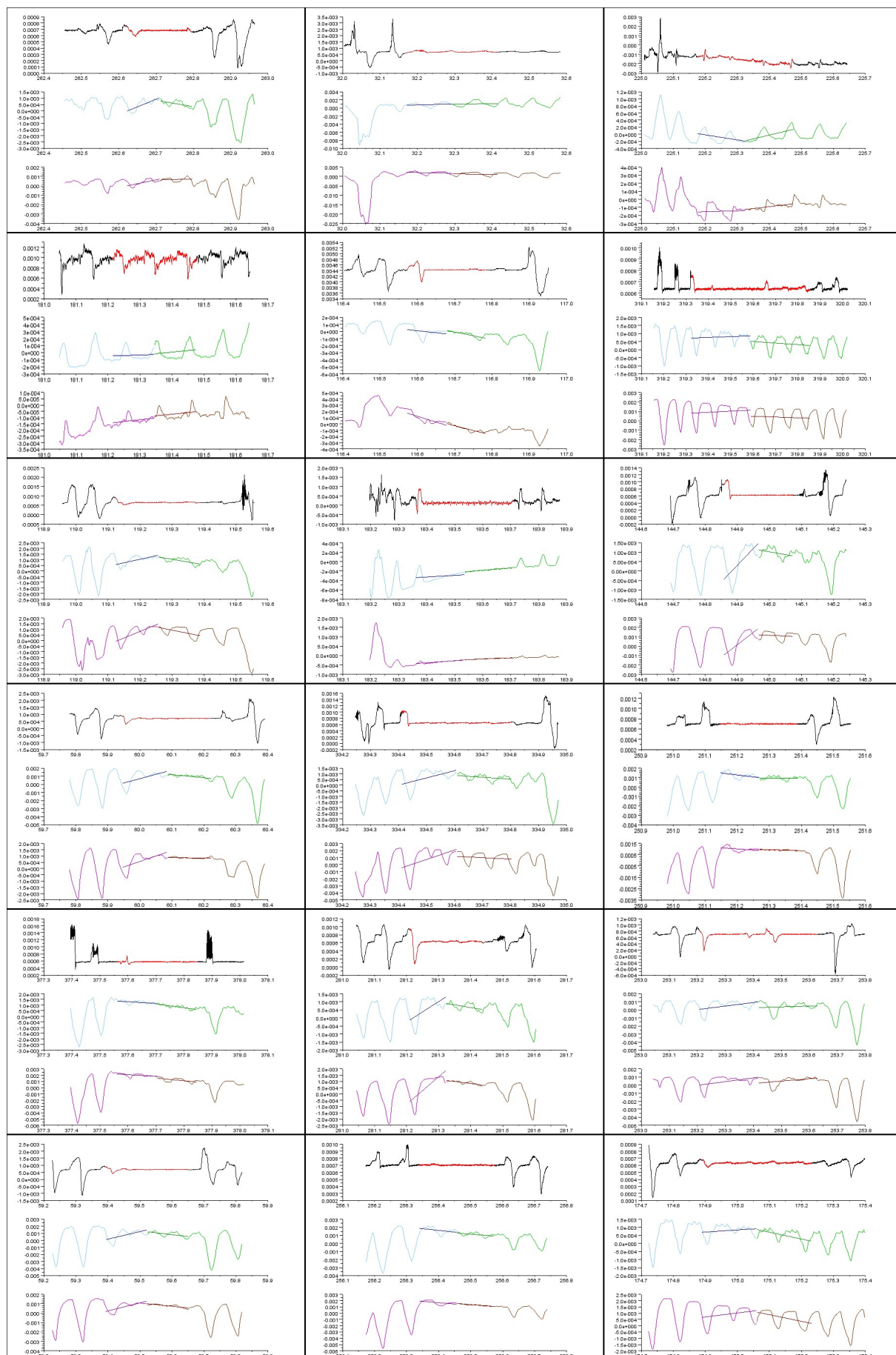


FIGURE 3.25 – Sélection d'hypnoses analysées comme des apnées par l'analyse automatique

Pour avoir une évaluation fiable des événements détectés par l'analyse automatique, il serait nécessaire de faire un relevé exhaustif de l'ensemble des événements analysés différemment par l'expert médical et de lui soumettre, pour qu'il les valide ou non. Cette évaluation est compliquée à réaliser et reste à effectuer.

### Reconnaissance du type des événements respiratoires

Notre algorithme de reconnaissance du type des événements respiratoires est assez efficace pour reconnaître les apnées obstructives et les apnées centrales, mais semble assez mauvais pour reconnaître les apnées mixtes.

Afin d'analyser les erreurs, nous avons généré une sélection d'apnées de chaque type pour lesquelles l'analyse automatique a conclu à un type différent :

- Sur la figure 3.26, on peut voir 6 apnées **obstructives** classées différemment par l'analyse automatique.
- Sur la figure 3.27, on peut voir 6 apnées **centrales** classées différemment par l'analyse automatique.
- Sur la figure 3.28, on peut voir 6 apnées **mixtes** classées différemment par l'analyse automatique.

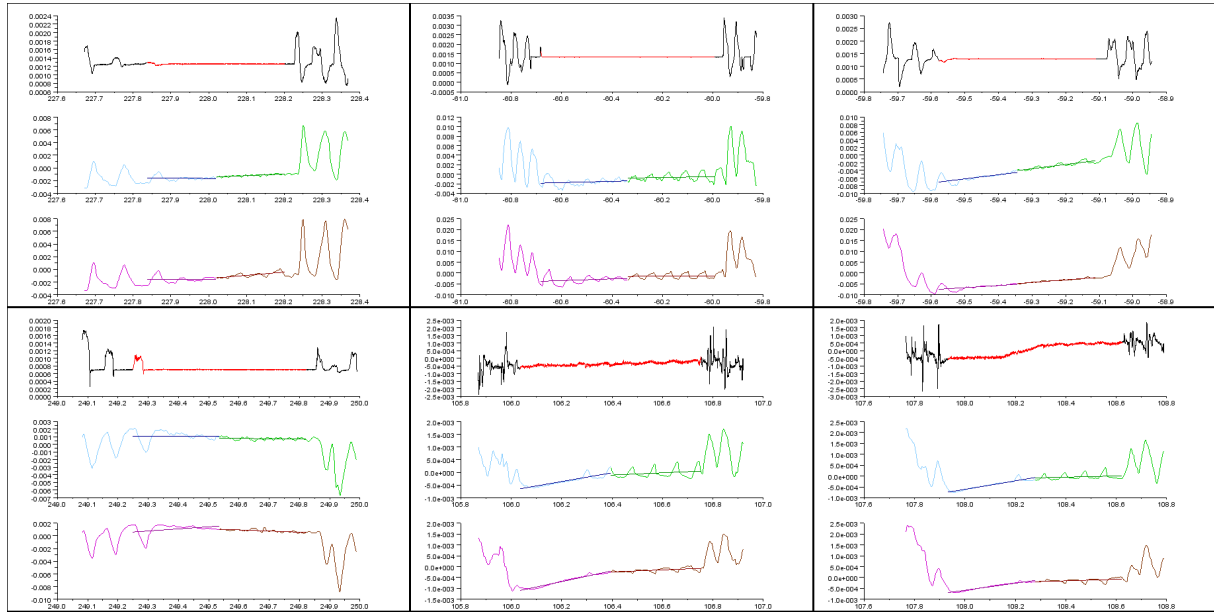


FIGURE 3.26 – Sélection d'apnées obstructives analysées différemment par l'analyse automatique

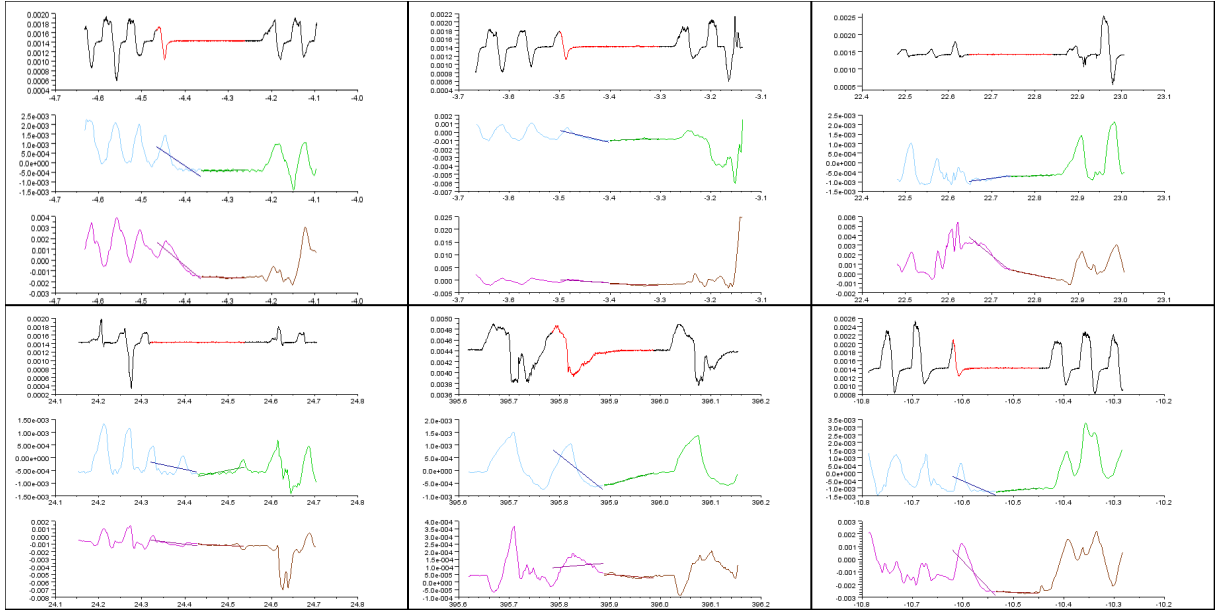


FIGURE 3.27 – Sélection d'apnées centrales analysées différemment par l'analyse automatique

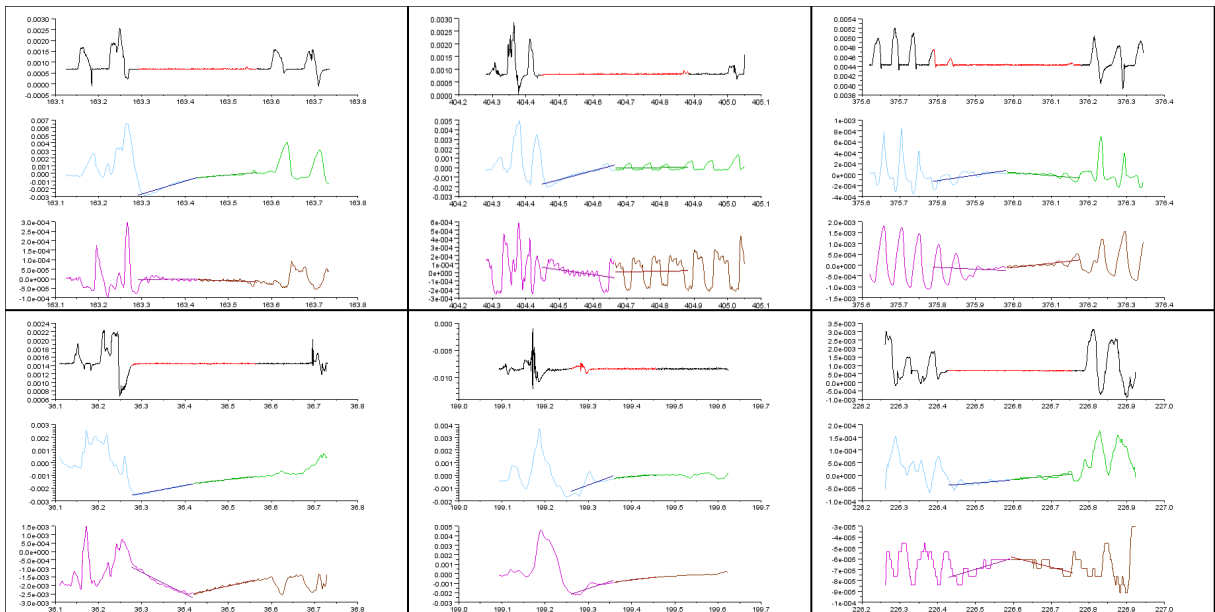


FIGURE 3.28 – Sélection d'apnées mixtes analysées différemment par l'analyse automatique

Nous pouvons tirer plusieurs conclusions au regard de ces courbes.

1. Il arrive que l'analyse de la présence ou absence d'effort respiratoire donne des résultats erronés car elle commence trop tôt. En effectuant l'analyse sur les 90% de signal situés en fin d'apnée, les résultats devraient être meilleurs, et répondre davantage à la définition donnée par l'AASM [31], qui demande que les critères d'amplitude soient respectés sur 90% de la durée de l'événement.
2. Il semble qu'il faille également revoir les seuils, et plus généralement les règles, pour décider s'il y a présence ou absence d'effort respiratoire, en particulier le seuil utilisé pour analyser la corrélation des mouvements thoraciques et abdominaux.

### 3 Aspects collaboratifs avec l'Université de Ballarat

Au cours de ce doctorat, j'ai eu la chance d'être accueilli pendant 5 semaines par le centre de recherche CIAO (Centre for Informatics and Applied Optimization) de l'Université de Ballarat dans l'État de Victoria en Australie. Ce centre de recherche dispose d'un laboratoire en Informatique médicale (Health Informatics Laboratory, HIL) créé à l'initiative du Dr Andrew STRANIERI. Intéressé pour travailler sur les données polysomnographiques de l'Hôpital Tenon avec différentes approches, cette expérience a été l'occasion de découvrir le fonctionnement d'un laboratoire étranger, de nouvelles approches face aux problématiques posées par l'automatisation de l'interprétation des données polysomnographiques. Le fruit de cette collaboration a été publié dans 3 papiers [56, 71, 79]

#### 3.1 Identification automatique de Stades de Sommeil

Il s'agit ici de tester plusieurs algorithmes pour identifier automatiquement les stades de sommeil. Quatre approches ont été testées :

1. réseau de neurones ;
2. prévision financière ;
3. optimisation non-lisse ;
4. Analyse du domaine de fréquences.

##### Réseau de neurones

Un réseau de neurones est composé de nœuds artificiels organisés en couches depuis la couche d'entrée vers la couche de sortie, en passant par une ou plusieurs couches intermédiaires (cachées). Les nœuds des couches intermédiaires peuvent s'activer en fonction de la valeur en entrée et ensuite transmettre aux nœuds connectés.

Les réseaux de neurones à décalage temporel sont un type de réseau de neurones dont la sortie à l'instant  $t_i$  dépend non seulement de l'entrée de l'instant  $t_i$  mais aussi des instants  $t_{i-1}$ ,  $t_{i-2}$ ,  $\dots$ ,  $t_{i-n}$  où  $n$  est le retard considéré par le réseau.

Il a été utilisé ici un réseau de neurones à décalage temporel avec une couche d'entrée, 3 couches cachées et 1 couche de sortie. La couche d'entrée comptait 6 nœuds (un par

voie polysomnographique utilisée pour le scorage des stades de sommeil, c'est-à-dire les 3 voies EEG, les 2 voies EOG et la voie EMG). Chaque couche cachée contenait 6 nœuds et la couche de sortie en comptait 5, un par stade de sommeil (Wakefulness, N1, N2, N3, REM). En entrée les signaux ont été normalisés sur l'intervalle  $[-1,1]$  puis découpés en séquences temporelles. Le réseau de neurones a été configuré avec un délai de 1 seconde (soit 100 vecteurs d'entrée). L'apprentissage a été effectué sur 500 époques. Au final, 76,15% des enregistrements ont été correctement classés.

### Prévision financière

Cette méthode a également pour hypothèse que la classification d'une séquence dépend de la classification des séquences précédentes. Elle utilise des probabilités conditionnelles sur les sous-séquences.

Les valeurs séquentielles sont converties en 5 classes d'intervalles : BI (Big Increase), SI (Small Increase), N (No change), SD (Small Decrease), BD (Big Decrease). Les seuils sont calculés en fonction des percentiles. L'hypothèse est qu'il existe des séquences de longueur maximale 7, sur l'une des voies polysomnographiques, qui peuvent servir de signature pour un stade de sommeil.

La matrice de confusion donnée dans le tableau 3.5 confronte les classifications faites par les experts à celles faites par l'analyse automatique par prévision financière.

TABLEAU 3.5 – Matrice de confusion de l'analyse automatique par Prévision Financière

		Analyse Automatique				
		W	N1	N2	N3	R
Expert	W	45%	23%	28%	1%	3%
	N1	4%	81%	12%	7%	13%
	N2	2%	9%	60%	7%	13%
	N3	0%	2%	2%	96%	0%
	R	4%	4%	20%	0%	71%

Ces résultats sont assez prometteurs et laissent entrevoir des améliorations possibles. L'ensemble d'apprentissage contenant une majorité d'époques de stade N2, il était plus difficile de trouver une signature pour les autres stades. Il faut donc une répartition plus équitable entre les différents stades de sommeil pour l'ensemble d'apprentissage.

Ensuite, les tests ont montré que l'algorithme était très sensible aux valeurs seuils utilisées pour les classes d'intervalles BI, SI, N, SD, BD. Il est donc à prévoir de travailler sur l'optimisation des valeurs de seuils.

Enfin, il a été constaté que les signatures sont des séquences courtes et qu'il n'est pas utile d'utiliser des séquences de taille 6 ou plus, car cela n'améliore pas vraiment les résultats, mais ralentit considérablement le temps de traitement.

### Optimisation non-lisse

L'identification des Stades de Sommeil par optimisation non lisse consiste à minimiser l'écart entre les courbes polysomnographiques et leur modélisation. Cette méthode procède à l'extraction de caractéristiques, comme il est pratiqué en traitement du signal, mais ces caractéristiques sont plus flexibles que les caractéristiques standards utilisées en traitement du signal et visent à décrire la forme des ondes. Ces caractéristiques peuvent être utilisées pour une description explicite des caractéristiques de la forme des ondes, comme un réseau de neurones, mais la dimension du problème est nettement moindre.

Les courbes EEG sont modélisées comme la somme de deux courbes sinusoïdales. La première courbe (de plus basse fréquence) représente la tendance générale sur l'ensemble de la période étudiée. La seconde (de plus haute fréquence) correspond au comportement local de la courbe autour de la tendance générale. L'amplitude de chaque courbe est modélisée par une fonction linéaire par morceaux. Cette approche permet d'obtenir des caractéristiques plus précises que celles utilisées dans le cas de sinusoïdes classiques où l'amplitude est constante. De plus, l'utilisation de fonctions linéaires par morceaux autorise à des changements abrupts dans les caractéristiques des ondes ; d'où le caractère non lisse de l'approche. Lors de l'implémentation, les expériences ont été conduites en utilisant les techniques d'optimisation de la bibliothèque GANSO (Ganso 2006 [7,9]).

Les expériences ont été conduites sur une courbe EEG, l'axe horizontal étant l'axe temporel. Dans un premier temps, la courbe sinusoïdale de plus haute fréquence est calculée (figure 3.29). Cette courbe est la première approximation de la voie EEG. En prenant en compte la tendance générale de la courbe, il est possible d'améliorer la précision de l'approximation. La figure 3.30 représente la tendance générale (en rouge) et une courbe noire calculée comme étant la différence entre la courbe originale et la première tendance. Finalement, la figure 3.31 représente la courbe finale, qui est une bonne approximation de la courbe originale. Lors des tests, la courbe a été découpée en période de 5 secondes ; chaque époque compte donc 6 « morceaux ». La dimension de ce problème est donc de 12 ( $2 \times 6$ ) ; elle serait de 500 pour un réseau de neurones. On peut donc envisager d'utiliser la sortie de ce problème comme une entrée d'un réseau de neurones.



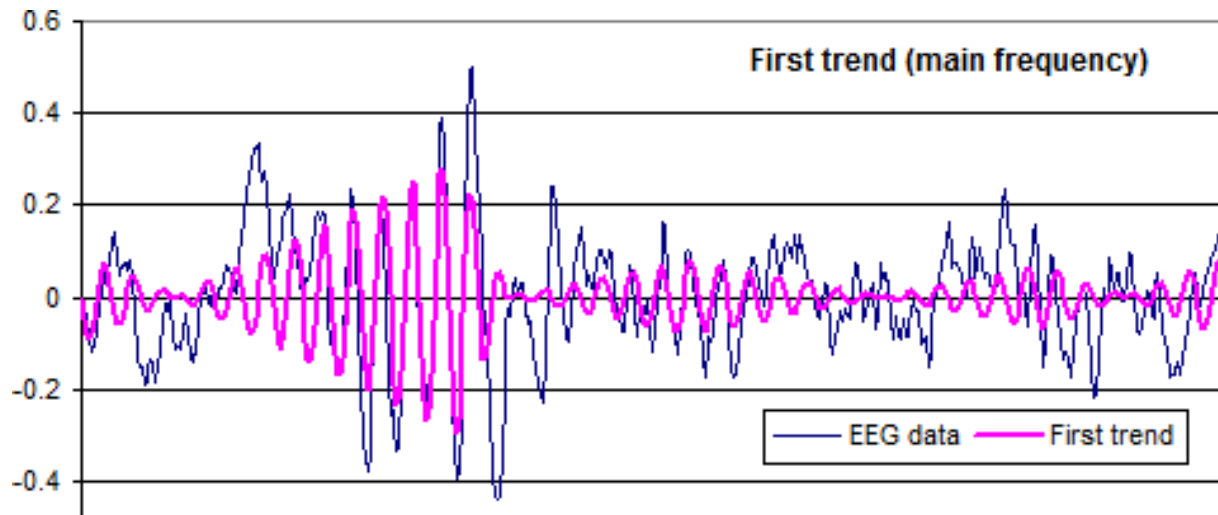


FIGURE 3.29 – Première tendance

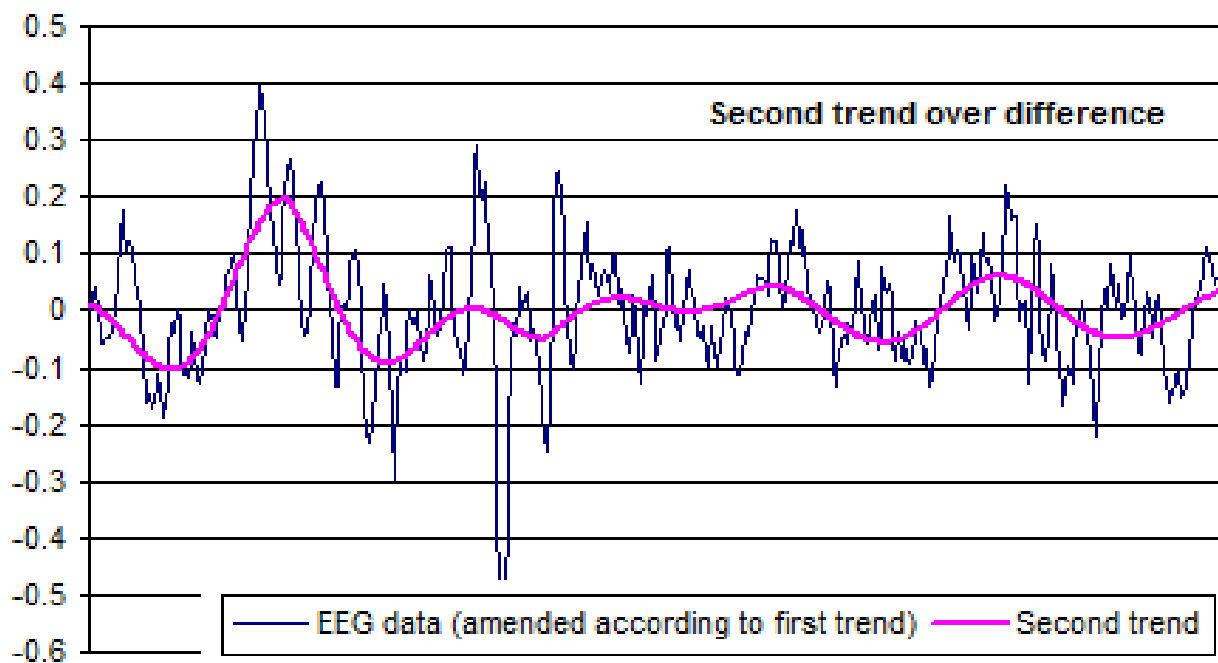


FIGURE 3.30 – Tendance générale

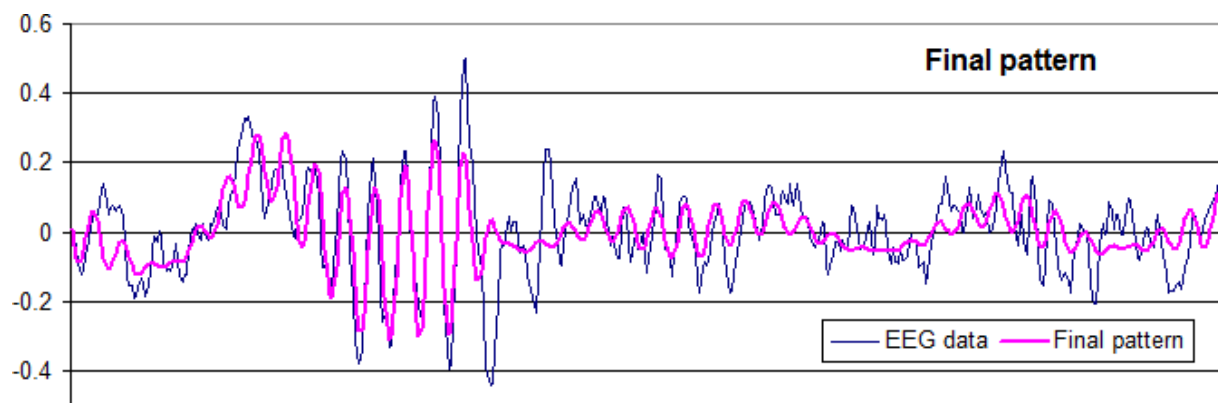


FIGURE 3.31 – Approximation Finale



### Analyse dans le domaine fréquentiel

L'approche par analyse en domaine fréquentiel est basée sur le principe de découper le signal en fenêtres et de transformer chacune d'entre elles dans le domaine fréquentiel, c'est ce qu'on appelle la *transformée de Fourier à court terme*. Le signal résultant est mis en correspondance avec un espace (ou une fonction) à deux dimensions : le temps et la fréquence. Les signaux EEG, EOG et EMG n'étant pas stationnaires, la conversion donne une précision limitée.

Les signaux EEG, EOG et EMG sont transformés dans le domaine fréquentiel par fenêtres de 30 secondes.

4 bandes de fréquences sont utilisées :

- $\delta$  :  $< 4Hz$  ;
- $\theta$  :  $4 - 7Hz$  ;
- $\alpha$  :  $8 - 13Hz$  ;
- $\beta$  :  $> 13Hz$ .

Une fois que les composantes fréquentielles sont séparées, on affiche la densité du spectre de fréquence par la formule suivante :

$$\Phi(\omega) = \frac{F(\omega) \cdot F^*(\omega)}{2\pi}$$

où  $F^*(\omega)$  est le nombre complexe conjugué de la matrice de fréquences.

Le sommeil paradoxal est généralement caractérisé par les caractéristiques suivantes :

- EEG de faible amplitude à fréquence rapide, ce qui se manifeste par une augmentation de  $\Phi(\beta)$  ;
- baisse relative des densités spectrales des composantes EEG de basse fréquence.

Ce sont ces règles qui sont utilisées par la reconnaissance du sommeil paradoxal.

La figure 3.32 montre les courbes des Transformées de Fourier à court terme pour les voies EEG, EOG, EMG mises en correspondance avec les stades de sommeil scorés manuellement par l'expert.

Le sommeil paradoxal commence par une hausse des densités spectrales des composantes EMG à très basses fréquences (EMG  $\Phi(\delta)$ ). Les positions 'a' et 'b' ne peuvent pas être considérées comme le début du sommeil paradoxal car les densités spectrales des composantes EEG à fréquences rapides (EMG  $\Phi(\alpha)$ ) augmentent rapidement en même temps que EMG  $\Phi(\delta)$ . 'c' n'est pas non plus le début du sommeil paradoxal car il devrait y avoir une petite montée de EEG  $\Phi(\beta)$ . 'd' est le début du sommeil paradoxal car il y a une hausse de EMG  $\Phi(\delta)$  et une légère hausse de EEG  $\Phi(\beta)$ . Il y a également une hausse pour toutes les fréquences pour l'EOG. Au point 'e', le sommeil paradoxal se termine par une hausse aigüe de EEG  $\Phi(\alpha)$ , signe du stade éveil. 'f' marque le début d'une nouvelle période de sommeil paradoxal car il y a une hausse de EMG  $\Phi(\delta)$  accompagnée par une petite hausse de EEG  $\Phi(\beta)$ . Cette période de sommeil paradoxal se termine au point 'g' car il y a une hausse de EEG  $\Phi(\alpha)$ . La dernière période de sommeil paradoxal, assez courte, n'a pas été détectée.

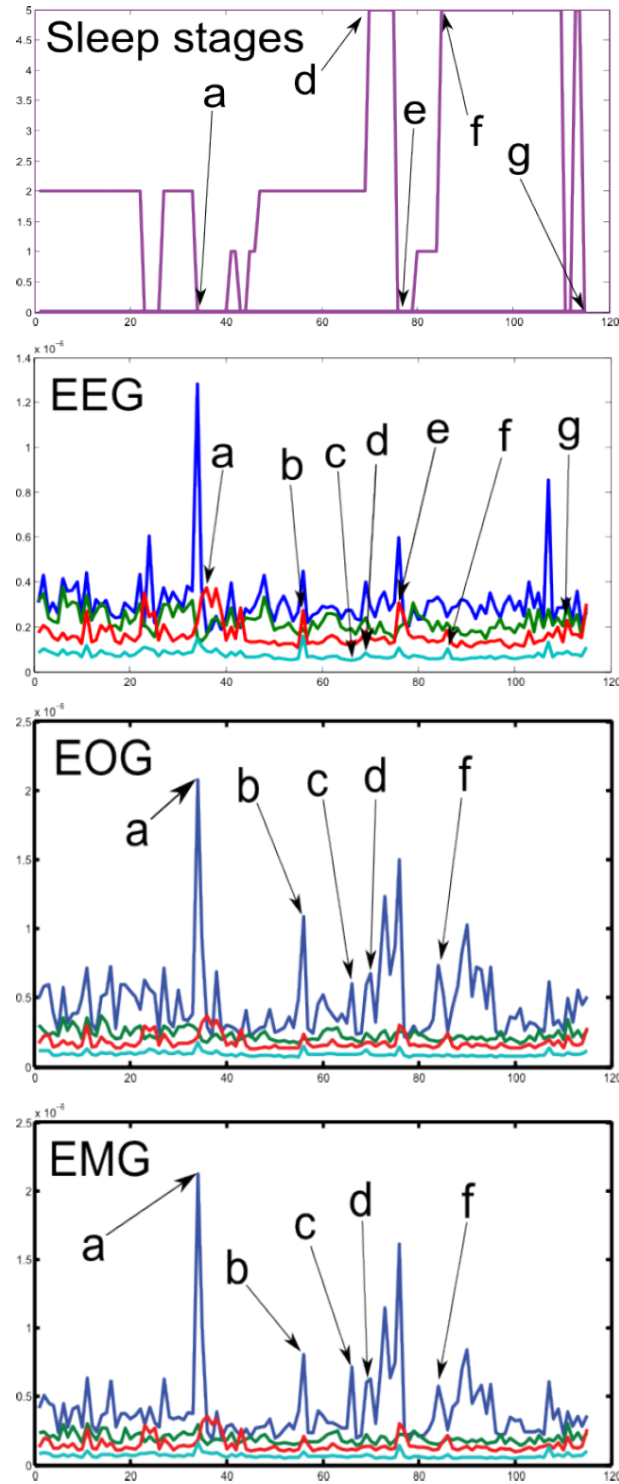


FIGURE 3.32 – Transformée de Fourier à court terme par bandes de fréquences des voies EEG, EOG et EMG et Hypnogramme

L'ensemble de ces règles qui sont inférées des caractéristiques du sommeil paradoxal, peuvent être implémentées en utilisant un automate. Toutefois, les seuils des composantes de fréquences différentes qui pourraient déclencher le changement d'état varient d'un cas à l'autre.

## Analyse et discussion

Ce travail permet de démontrer que le processus d'interprétation automatique des stades de sommeil est complexe et nécessite le recours à diverses méthodes. Cette diversité permet de surmonter le problème de la « traduction » des règles du manuel de l'AASM [31] en algorithmes.

Les meilleurs résultats sont ceux obtenus par le réseau de neurones à décalage temporel, ce qui permet de gérer une importante quantité de données. La complexité reste élevée et on peut imaginer qu'elle serait réduite par l'utilisation préliminaire d'une optimisation non-lisse et permettrait d'avoir un nombre de paramètres en entrée beaucoup moins élevé. Il peut également être envisagé d'utiliser les points forts de chaque algorithme et de les combiner. Ainsi, le stade REM serait détecté par l'analyse dans le domaine fréquentiel ; le stade N3 serait détecté par « prévision financière ». Le stade N2 s'est avéré être un stade difficile à reconnaître, ce qui s'explique par la présence spécifique de patterns courts que les méthodes présentées ici ne peuvent pas reconnaître avec une fiabilité suffisante. Il est envisageable de tester des méthodes de reconnaissance en utilisant l'optimisation non lisse.

Il serait alors envisagé d'utiliser un système de reconnaissance des stades de sommeil en plusieurs étapes :

1. modélisation du signal par optimisation non lisse ;
2. réseau de neurones ;
3. affinage des résultats, notamment pour certains stades de sommeil, en utilisant l'analyse en domaine fréquentiel ou la prévision financière.

Ce travail a été publié dans [71].

## 3.2 Détection des complexes K par optimisation non lisse

L'application de l'optimisation non lisse à la détection des complexes K consiste à rechercher un modèle de courbes tel que l'écart à la courbe EEG modélisée (somme des écarts absolus) soit minimal. Cette méthode consiste à extraire des caractéristiques des courbes pour obtenir une description de la forme des ondes, et ainsi réduire la dimension des problèmes de classification. La courbe EEG est modélisée sous la forme de la somme de 2 courbes sinusoïdales. L'amplitude de chaque courbe est modélisée par une fonction linéaire par morceaux. Ceci permet d'obtenir des éléments de courbe plus proches de la courbe modélisée que ce qui serait obtenu par l'utilisation de fonctions sinus avec une amplitude constante.

Il s'agit ici de considérer les définitions suivantes : Les *splines polynomiales* sont des fonctions polynomiales par morceaux. Dans la plupart des applications, les *splines polynomiales* sont continues. Les points où les polynômes sont joints ensemble sont appelés les *nœuds* des splines. Les *nœuds* des splines peuvent être fixes ou libres. Le plus haut degré des polynômes constituant la spline est appelé le *degré* de la spline.

Dans cette étude, l'amplitude du signal est modélisée par une fonction spline, préférée à une constante. Cette approche permet plus de flexibilité car l'amplitude peut varier sur l'intervalle de 10 secondes.

Considérons la fonction spline  $S_m$  suivante :

$$S_m(A, \theta, t) := a_0 + \sum_{j=1}^m a_{1j} \cdot t^j + \sum_{i=2}^n \sum_{j=1}^m a_{ij} ((t - \theta_{i-1})_+)^j$$

où :

- m est le degré de la spline ;
- $\theta = (\theta_1, \dots, \theta_n)$  sont les nœuds. On utilisera 1 à 3 nœuds dans ce projet ;
- 

$$(\xi(x))_+ := \begin{cases} \xi(x), & \xi(x) > 0 \\ 0, & \xi(x) \leq 0 \end{cases}$$

- $A = (a_0, a_{11}, \dots, a_{nm}) \in \mathbb{R}^{mn+1}$  est appelé le vecteur des paramètres de la spline.

Selon le problème, les nœuds peuvent être fixes ou libres.

Afin d'utiliser des splines polynomiales pour l'amplitude, il est nécessaire de connaître les paramètres et les nœuds de la spline. Ceux-ci peuvent être obtenus en résolvant le problème d'optimisation suivant :

$$\min_{X, \omega, \tau} \sum_{i=1}^N |y_i - \text{Amp}(X, t_i) \cdot \sin(2\pi\omega t_i + \tau)|$$

où

- $y_i, i=1, \dots, B$  est la valeur de l'enregistrement EEG à l'instant  $t_i$  ;
- $\text{Amp}(X, t)$  est la fonction Amplitude, modélisée par une spline polynomiale ;
- $\omega$  est la fréquence ;
- $\tau$  est la valeur qui permet que les courbes soient en phase ;
- $X$  est le vecteur qui caractérise la fonction amplitude.

Pour résoudre le problème précédent, il est nécessaire d'utiliser des méthodes d'optimisation non lisse. Les problèmes d'optimisation non lisse sont généralement difficiles et longs à résoudre. Pour éviter ce problème, il est possible d'utiliser la méthode des moindres carrés plutôt que la somme des écarts absolus. Cependant, la méthode des moindres carrés perd de sa robustesse quand il y a beaucoup d'outliers, ce qui est le cas dans les courbes EEG enregistrées ici. Il est donc préférable d'utiliser la somme des écarts absolus pour extraire les caractéristiques des courbes. Nous allons ici procéder ainsi pour détecter les complexes K.

Nous allons utiliser les modèles suivants pour la fonction amplitude  $\text{Amp}$  :

- **Spline linéaire à nœuds fixes avec 2 intervalles ;**

$$\text{Amp} = \text{Amp}_1(X, t) = a_0 + a_1 \cdot t + a_2 \cdot (t - \theta_1)_+$$

Le nœud  $\theta_1$  correspond au centre de l'intervalle de 10 secondes (situé donc à 5 secondes du début de chaque période) ;  $X = (a_0, a_1, a_2)$

Ce modèle utilisé pour la fonction d'amplitude est pertinent pour les époques pour lesquelles il y a un changement d'amplitude approximativement en leur milieu.

Par exemple, si l'amplitude augmente dans la première moitié et reste constante sur la deuxième moitié, on peut attendre  $a_1 > 0$  et  $a_2 = -a_1$ . Ce modèle convient également aux périodes pour lesquelles l'amplitude ne varie pas. Par exemple, si l'amplitude décroît sur toute la durée de la période, on peut attendre  $a_1 < 0$  et  $a_2 = 0$ . Il s'agit d'un problème d'optimisation de dimension 2.

– **Spline linéaire à nœuds libres avec 2 intervalles ;**

$$Amp = Amp_2(X, T) = a_0 + a_1 \cdot t + a_2 \cdot (t - \theta_1)_+$$

$Amp_2$  est similaire à  $Amp_1$ , mais  $\theta_1$  est cette fois considérée comme une variable dans le problème d'optimisation. On a alors  $X = (a_0, a_1, a_2, \theta_1)$ .  $Amp_2$  est donc plus flexible que  $Amp_1$ , mais le problème d'optimisation associé est plus compliqué, ce qui signifie que le problème est algorithmiquement plus complexe et que la solution peut ne pas être parfaitement satisfaisante.

Ce modèle est pertinent pour tous les cas décrits précédemment, mais offre plus de liberté pour la position du changement d'amplitude car celle-ci peut être n'importe où, et non obligatoirement au milieu. Il s'agit donc d'un problème d'optimisation de dimension 6.

– **Amplitude en pic aigu ;**

$$Amp = Amp_3(X, t) = \max \{a_0, -a_1 \cdot |t_i - a_2| + a_3\} \quad a_j \geq 0, j = 0, \dots, 3, a_2 \leq 10$$

Ce modèle est pertinent pour toutes les périodes pendant lesquelles l'amplitude est relativement constante pendant la période ( $a_1 = 0$ ) ou s'il y a une hausse soudaine et pointue de l'amplitude alors qu'elle est à un même niveau sur tout le reste de la période. On a alors  $X = (a_0, a_1, a_2, a_3)$ . Il s'agit donc d'un problème d'optimisation de dimension 6.

Ce modèle est un cas particulier de la spline linéaire avec des conditions restrictives sur les paramètres de la spline.

Les tests ont été faits en utilisant la bibliothèque GANSO [7, 9], qui implémente plusieurs méthodes d'optimisation globale, non-lisse, non-convexe et non linéaire. En pratique, il n'y a pas de méthode d'optimisation capable de trouver un minimum global à n'importe quelle fonction, en particulier si cette fonction est non-convexe et non-lisse, ce qui est le cas dans notre problème énoncé ci-dessus. Nous pouvons seulement affirmer que la solution est un optimum local (optimum dans un certain voisinage) et, en conséquence, en pratique, une solution à un problème plus simple peut être meilleure que celle d'un problème plus compliqué, même si celui-ci, en théorie, décrit plus précisément le comportement de l'amplitude. Cette règle s'applique d'autant plus que la complexité du problème d'optimisation augmente.

Il s'est avéré suite aux expériences avec  $Amp_1$ ,  $Amp_2$  et  $Amp_3$  que  $Amp_1$  est la meilleure solution. En effet,  $Amp_1$  est plus rapide à obtenir que  $Amp_2$ , et les nœuds  $\theta$  de  $Amp_2$  sont situés au tout début ou à la toute fin de la période quasi systématiquement, et sont ainsi inutiles.  $Amp_3$  a échoué à découvrir toute hausse soudaine de l'amplitude. C'est pourquoi nous nous intéresserons exclusivement aux résultats associés à  $Amp_1$ .

Nous remarquons que la solution cherchée semble nécessiter de modéliser aussi la fréquence du signal par une fonction polynomiale. En effet, celle-ci décroît significativement lors de la survenue du complexe K. Cependant, cela augmente la dimension du problème d'optimisation sans en améliorer significativement les résultats.

Les tests ont consisté à deux étapes successives :

1. Dans un premier temps, l'erreur a été minimisée.

$$\min_{X, \omega, \tau} \sum_{i=1}^N |f(t_i) - \text{Amp}(X, t_i) \cdot \sin(2\pi\omega t_i + \tau)|$$

Supposons que  $(X_1, \omega_1, \tau_1)$  est la solution obtenue après la première étape, alors la courbe associée aura pour équation  $W_1 = \text{Amp}(X_1, t) \cdot \sin(2\pi\omega_1 t + \tau_1)$ .

2. Dans la seconde étape, nous minimisons l'erreur sur un nouvel ensemble de données, créé en substituant les données initiales par la différence entre les données initiales et la courbe obtenue lors de la première étape.

$$\min_{X, \omega, \tau} \sum_{i=1}^N |y_i - W_1 - \text{Amp}(X, t_i) \cdot \sin(2\pi\omega t_i + \tau)|$$

Supposons maintenant que  $(X_2, \omega_2, \tau_2)$  est la solution obtenue à la seconde étape, la courbe aura alors pour équation  $W_2 = \text{Amp}(X_2, t_i) \cdot \sin(2\pi\omega_2 t_i + \tau_2)$ .

En conséquence, nous extrayons deux tendances  $W_1$  et  $W_2$  ; leur somme est le modèle de la courbe EEG sur la période de 10 secondes. Ensuite, les paramètres de cette tendance (de dimension double par rapport au problème d'optimisation correspondant) forment la sortie du problème d'optimisation non lisse. Dans le cas de  $\text{Amp}_1$ , la sortie est de dimension 10. Nous prenons aussi en compte trois autres paramètres qui caractérisent l'amélioration de la fonction recherchée après la résolution du problème d'optimisation non lisse. Nous avons donc réduit, par cette modélisation, un problème initial de dimension 1000 à 13 caractéristiques essentielles.

Il faut bien noter que la solution obtenue par la résolution de ce problème en deux étapes est différente de celle qui aurait été obtenue par une seule étape. Ce choix permet de réduire la dimension du problème et la complexité des algorithmes.

Les résultats ont ensuite été injectés dans un algorithme d'apprentissage, en utilisant la plateforme Weka. Plusieurs algorithmes ont été testés.

Dans un premier temps, les algorithmes ont été testés sur les données brutes. L'exactitude (*Accuracy*) est calculée par la formule suivante :

$$\frac{VN + VP}{VN + VP + FP + FN}$$

Les résultats sont présentés dans le tableau 3.6.

TABLEAU 3.6 – Résultats obtenus sous Weka sur les données brutes (sans modélisation)

Algorithme	Exactitude sur l'ensemble de test
RBF	73,97%
Random Forest	63,01%
J48	56,16%
Lazy IB1	56,16%
J48graft	53,42%
Random Tree	53,42%
Lazy IB5	49,32%
Lazy IB10	47,95%
LibSVM	47,95%
LMT	45,21%
SMO	43,84%
Logistic	39,73%
MLP	ECHEC

D'autres tests ont ensuite été effectués sur la plateforme Weka après un preprocessing, c'est-à-dire après la modélisation de la courbe par une spline polynomiale par morceaux, grâce aux méthodes de la bibliothèque GANSO [7, 9] (ECAM, DFBM, DFBMECAM, DSO). ECAM0.1 et ECAM0.001 sont des modifications de ECAM avec une modification du paramètre Lipschitz. Par défaut (dans ECAM), ce paramètre vaut 1 ; il vaut 0,1 dans ECAM0.1 et 0,001 dans ECAM0.001. Ceci permet notamment d'accélérer les temps de traitement de l'algorithme. La courbe est donc tout d'abord modélisée sous GANSO avec ECAM0.001, ECAM0.1, ECAM, DFBM, DFBMECAM, DSO. Ensuite, le résultat est injecté dans l'un des algorithmes d'apprentissage de Weka (LibSVM, Logistic, MLP, RBF, SMO, Lazy IB1, Lazy IB5, Lazy IB10, J48, J48graft, LMT, Random Forest, Random Tree). L'exactitude obtenue pour la reconnaissance des complexes K sur l'ensemble de tests est affichée pour chaque couple d'algorithmes (GANSO / WEKA) dans le tableau 3.7.

TABLEAU 3.7 – Résultats obtenus sous Weka sur les données après modélisation (preprocessing)

Algorithme	ECAM0.001	ECAM0.1	ECAM	DFBM	DFBMECAM	DSO
LibSVM	52,54%	62,71%	59,32%	54,24%	54,24%	54,24%
Logistic	54,23%	54,23%	55,93%	59,32%	66,01%	55,93%
MLP	57,62%	45,76%	55,93%	59,32%	57,63%	61,02%
RBF	62,71%	44,02%	61,02%	61,02%	57,63%	55,93%
SMO	45,76%	50,85%	64,41%	62,71%	55,93%	47,46%
Lazy IB1	62,72%	50,85%	54,24%	40,68%	61,02%	64,41%
Lazy IB5	64,41%	62,72%	62,72%	50,85%	64,41%	62,72%
Lazy IB10	54,24%	54,24%	61,02%	59,32%	49,15%	59,32%
J48	49,15%	54,24%	50,85%	55,93%	55,93%	55,93%
J48graft	49,16%	54,24%	50,85%	55,93%	55,93%	55,93%
LMT	47,46%	55,93%	62,71%	61,02%	62,72%	45,77%
Random Forest	62,72%	55,93%	57,62%	55,93%	55,93%	55,93%
Random Tree	62,71%	54,24%	54,24%	44,07%	59,32%	66,10%

Nous remarquons ici qu'aucun des algorithmes de Weka n'a échoué, ce qui s'explique par la dimension bien plus faible du problème. Les résultats sont également nettement améliorés par rapport aux premiers tests avec la modélisation, sauf pour RBF et Random Forest. Les meilleurs résultats sont obtenus par l'association de DSO et de Random Tree (66,1%) ou par l'association de DFBMECAM et Logistic (66,01%).

Il semble cependant que les résultats restent moins bons que ceux de RBF appliqués aux données brutes. Ce n'est pas tout à fait le cas car il y a un grand nombre de faux négatifs dans le résultat de RBF, comme le montre la matrice de confusion :

$$\begin{pmatrix} VP & FN \\ FP & VN \end{pmatrix} = \begin{pmatrix} 20 & 15 \\ 4 & 34 \end{pmatrix}$$

Un tel taux de faux négatifs n'est pas acceptable car un tel support au diagnostic se doit de préférer présenter un ensemble de cas « suspectés » de correspondre à des complexes K ; l'expert invalidera ensuite lui-même les éventuels faux positifs. Il est donc important que l'ensemble des « vrais » complexes K soient détectés, et qu'il y ait un nombre quasi nul de faux négatifs.

Nous proposons donc une procédure en plusieurs étapes :

1. Appliquer l'algorithme RBF sur les données brutes.
2. Appliquer la modélisation puis la classification sur les segments n'ayant pas été détectés comme complexes K lors de la première étape.

Cette méthode ne permet pas d'améliorer l'exactitude, mais réduit considérablement le nombre de faux négatifs. Ce but est atteint de façon optimale en utilisant le couple (ECAM0.1 / MLP) lors de la deuxième étape. L'exactitude reste ainsi à 74%. La précision vaut 73,5% (au lieu de 83,3%) ; le rappel vaut 71,4% (au lieu de 57,1%). La F-mesure vaut alors 72,5% (au lieu de 67,8%), ce qui est un résultat satisfaisant, car on augmente effectivement le rappel (de 14,3%), ce qui était le but, sans perdre trop de précision.

Ce travail a été publié dans [56].

### 3.3 Détection des complexes K utilisant un apprentissage hybride-synergique

Il s'agit ici de reconnaître des complexes K. L'ensemble de données mis à notre disposition par l'Hôpital Tenon est découpé sous la forme de segments EEG de 10 secondes, chacun étant étiqueté comme contenant un complexe K ou non. C'est donc un apprentissage non standard, car on ne travaille pas directement sur des instances à identifier, mais sur des *sacs* d'instances parmi lesquelles certaines peuvent correspondre à la forme recherchée. Les instances ne sont pas étiquetées individuellement ; seul le *sac* l'est. Un tel problème d'apprentissage appartient donc à une famille appelée *apprentissage à instances multiples* ou *multi-instance learning (MIL)*. Seront considérées comme instances d'un *sac* (ou segment) donné l'ensemble des sous-segments du segment étiqueté. Pour un segment étiqueté positivement, il existe au moins un sous-segment qui est un complexe K.



Dans un premier temps, afin de nettoyer les courbes du bruit, un lissage *triangulaire* par fenêtres mobiles (à coefficients non identiques) est appliqué. Le résultat de ce lissage est montré sur la figure 3.33.



FIGURE 3.33 – Lissage triangulaire utilisé en prétraitement des données

Une des difficultés qui se pose à la reconnaissance des complexes K est que leur définition est imprécise. Selon l'AASM [31], le complexe K est défini comme *une onde négative aux contours bien délimités suivie immédiatement par une onde positive se détachant des ondes de fond, d'une durée totale de plus d'une demie-seconde, habituellement maximale en amplitude sur les dérivation frontales*. Cette définition, bien que floue pour suffire dans l'utilisation d'algorithmes standards de traitement du signal, donne cependant deux informations utilisables pour un apprentissage : le complexe K dure plus de 0,5 seconde et il se détache du fond, notamment en amplitude.

Il est nécessaire de définir sur quelles instances sera appliqué l'algorithme d'apprentissage. Nous utiliserons ici des segments de 6 secondes avec un pas de 0,1 seconde entre 2 instances. La durée de 6 secondes semble convenir car les complexes K durent entre 0,5 seconde et 1,5 seconde ; un segment de 6 secondes convient pour détecter un événement de 1,5 seconde prenant en compte l'environnement afin de pouvoir différencier la forme recherchée des ondes de fond. Au total, le segment de 10 secondes contiendra donc 41 instances, c'est-à-dire 41 sous-segments, qui seront analysés par l'algorithme d'apprentissage. Les segments  $B_i$  peuvent donc être représentés de la façon suivante :  $B_i = \{I_{i1}, I_{i2}, \dots, I_{im}\}$  avec  $m=41$ . À chaque segment  $B_i$  est associée la valeur 1 si l'expert a indiqué qu'il contenait un complexe K, 0 sinon

La figure 3.34 illustre les segments qui seront analysés pour un segment donné et étiqueté par l'expert.

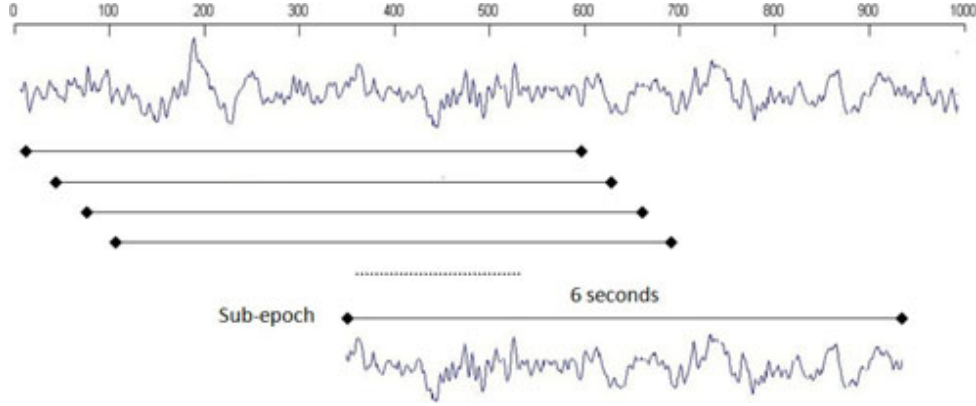


FIGURE 3.34 – Division des segments EEG en instances de 6 secondes

Dans un premier temps, le sous-segment est découpé en 3 parties de 2 secondes chacune. L'objectif de ce découpage est de déterminer laquelle des 3 parties contient, avec le maximum de vraisemblance, le complexe K. Cette partie sera appelée  $S_A$ , les deux autres seront appelées  $S_B$  et  $S_C$ .

Chaque instance sera ensuite décrite par un vecteur de paramètres extraits du signal analysé.

Afin de comprendre ces différents paramètres, il convient de définir les variables suivantes :

- $X_{max}$  est la valeur du maximum d'amplitude sur la partie  $S_A$ .
- $X_{min}$  est la valeur du minimum d'amplitude sur la partie  $S_A$ .
- $P_{max}$  est l'abscisse du point  $X_{max}$ .
- $P_{min}$  est l'abscisse du point  $X_{min}$ .
- $P_{start}$  est l'abscisse du premier point où la courbe croise l'axe médian quand on part de  $P_{max}$  et qu'on s'éloigne de  $P_{min}$ .
- $P_{end}$  est l'abscisse du premier point où la courbe croise l'axe médian quand on part de  $P_{min}$  et qu'on s'éloigne de  $P_{max}$ .
- $P_{middle}$  est l'abscisse du premier point où la courbe croise l'axe médian quand on part de  $P_{max}$  et qu'on va vers  $P_{min}$ .

Ces différents points sont représentés sur la figure 3.35.

Les lignes horizontales  $t_1, t_2, t_3, t_4$  ont pour valeur :

- $t_1 = \frac{2 \cdot X_{max}}{3}$ .
- $t_2 = \frac{X_{max}}{3}$ .
- $t_3 = \frac{X_{min}}{3}$ .
- $t_4 = \frac{2 \cdot X_{min}}{3}$ .

La partie  $S_A$  sera décrite par un ensemble de 12 paramètres décrits dans le tableau 3.8. Les deux autres parties,  $S_B$  et  $S_C$ , considérées comme contenant des ondes de fond, comparativement au complexe K potentiel de la partie  $S_A$ , seront décrites par 10

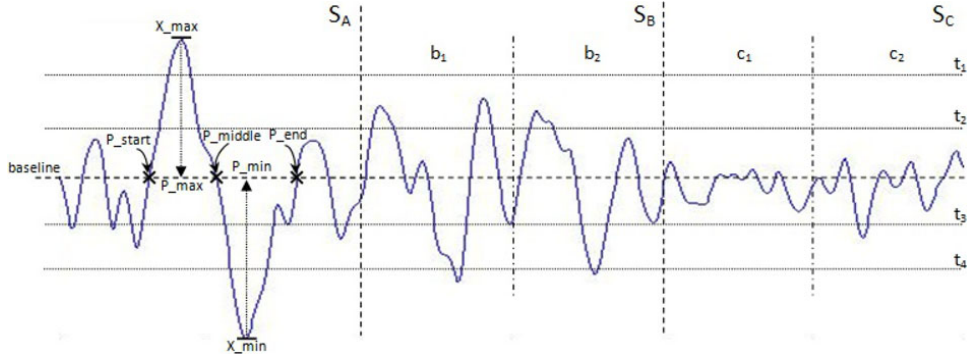


FIGURE 3.35 – Définition des paramètres sur le segment « instance » de 6 secondes

paramètres décrits dans le tableau 3.9. Enfin, le tableau 3.10 donne un ensemble de 7 paramètres utiles pour analyser les différences entre la partie  $S_A$  d'une part et les parties  $S_B$  et  $S_C$  d'autre part.

 TABLEAU 3.8 – Paramètres extraits du segment  $S_A$ 

Paramètre	Valeur	Description
$f_1$	$X_{max}$	Amplitude maximale
$f_2$	$X_{min}$	Amplitude minimale
$f_3$	$f_1 - f_2$	Différence d'amplitude entre l'amplitude maximale et l'amplitude minimale
$f_4$	$ P_{max} - P_{min} $	Nombre de points enregistrés entre $P_{max}$ et $P_{min}$
$f_5$	$f_3/f_4$	Acuité de la courbe entre le point max et le point min
$f_6$	$ f_2  /  f_1 $	Rapport entre l'amplitude maximale et l'amplitude minimale
$f_7$	$R_{max,min}$	Taux de référence de croisement de l'axe médian entre $P_{max}$ et $P_{min}$
$f_8$	$ P_{start} - P_{min} $	Durée du complexe K potentiel
$f_9$	$ P_{middle} - P_{start} $	Durée de la partie positive du complexe K potentiel
$f_{10}$	$ P_{middle} - P_{end} $	Durée de la partie négative du complexe K potentiel
$f_{11}$	$ f_1/f_9 $	Acuité de la partie positive du complexe K potentiel
$f_{12}$	$ f_2/f_{10} $	Acuité de la partie négative du complexe K potentiel

 TABLEAU 3.9 – Paramètres extraits des segments  $S_B$  et  $S_C$  (signal de fond)

Paramètre	Valeur	Description
$f_{13}$	$max_B$	Amplitude maximale sur l'intervalle $S_B$
$f_{14}$	$max_C$	Amplitude maximale sur l'intervalle $S_C$
$f_{15}$	$min_B$	Amplitude minimale sur l'intervalle $S_B$
$f_{16}$	$min_C$	Amplitude minimale sur l'intervalle $S_C$
$f_{17}$	$f_{13} - f_{15}$	Différence d'amplitude entre le min et le max sur l'intervalle $S_B$
$f_{18}$	$f_{14} - f_{16}$	Différence d'amplitude entre le min et le max sur l'intervalle $S_C$
$f_{19}$	$f_{17}/f_3$	Importance de l'amplitude relative sur l'intervalle $S_B$
$f_{20}$	$f_{18}/f_3$	Importance de l'amplitude relative sur l'intervalle $S_C$
$f_{21}$	$\frac{max_{b1} + max_{b2} + max_{c1} + max_{c2}}{4}$	Niveau d'amplitude positive des intervalles <i>de fond</i>
$f_{22}$	$\frac{min_{b1} + min_{b2} + min_{c1} + min_{c2}}{4}$	Niveau d'amplitude négative des intervalles <i>de fond</i>

L'objectif attendu par la communauté médicale est qu'un outil de support à la détection des complexes K surligne, sur les courbes, les potentiels complexes K détectés.

TABLEAU 3.10 – Paramètres extraits des segments  $S_B$  et  $S_C$  (signal de fond)

Paramètre	Valeur	Description
$f_{23}$	$f_1 - f_{21}$	Niveau de différence d'amplitude positive entre l'intervalle $S_A$ et les intervalles de fond
$f_{24}$	$f_{22} - f_2$	Niveau de différence d'amplitude négative entre l'intervalle $S_A$ et les intervalles de fond
$f_{25}$	$\frac{f_{21} - f_{22}}{f_3}$	Distinguabilité entre l'intervalle $S_A$ et les intervalles de fond
$f_{26}$	$R_{t_1}$	Nombre de points du segment croisant l'axe $t_1$
$f_{27}$	$R_{t_2}$	Nombre de points du segment croisant l'axe $t_2$
$f_{28}$	$R_{t_3}$	Nombre de points du segment croisant l'axe $t_3$
$f_{29}$	$R_{t_4}$	Nombre de points du segment croisant l'axe $t_4$

Un algorithme qui se contenterait d'identifier des segments de 10 secondes comme contenant un complexe K ne satisferait pas un médecin. C'est pourquoi il faut appliquer un algorithme qui va, non seulement étiqueter les segments positifs, mais aussi en extraire les instances correspondant le plus vraisemblablement à la forme recherchée. Ceci est possible en utilisant une machine à vecteurs de support (*Support Vector Machine (SVM)*).

De plus, il est nécessaire d'équilibrer les tailles des classes en vue de l'apprentissage. En effet, nous disposons de 54 (17%) segments étiquetés positivement, et 272 (83%) segments étiquetés négativement. Il est possible d'augmenter le nombre d'instances d'une classe minoritaire en utilisant la méthode SMOTE (*Synthetic Minority Oversampling Technique*) [14]. Nous utiliserons ici un algorithme de classification d'instances représentatives. Cet algorithme fonctionne en 2 étapes. La première sert à identifier des instances représentatives extraites depuis les segments et à les collecter dans un nouvel ensemble de données. La seconde consiste à élaborer un algorithme d'apprentissage sur ce nouvel ensemble de données.

### Identification d'instances représentatives

Soit  $\mathcal{D} = \{X, Y\}$  l'ensemble d'apprentissage.  $X$  contient  $n$  segments :  $X = \{B_1, B_2, \dots, B_n\}$ .  $Y$  est l'ensemble des étiquettes :  $Y = \{L_1, L_2, \dots, L_n\}$ . Chaque segment  $B_i$  contient  $m$  instances.  $B_i = \{I_{i1}, I_{i2}, \dots, I_{im}\}$  et son étiquette est  $L_i$ . Soit  $\delta_{pos}$  et  $\delta_{neg}$  les seuils d'un réseau de neurones pour les segments *positifs* et *négatifs* respectivement. Ici,  $\delta_{pos} = 1$  et  $\delta_{neg} = 0$ .

L'algorithme d'identification d'instances représentatives utilise un réseau de neurones sur les instances ; les paramètres sont ajustés en fonction de l'erreur globale :

$$Err = \frac{1}{2} \times \sum_{i=1}^n (\max_{1 \leq j \leq m_{ij}} o_{ij} - \delta_i)^2$$

où :

- $o_{ij}$  est la sortie du réseau de neurones.
- $\delta_{ij}$  vaut  $\delta_{pos}$  ou  $\delta_{neg}$ , en fonction de la valeur de  $L_i$  l'étiquette du segment  $B_i$ .

Selon cette définition, l'erreur vaudra 0 si les 2 conditions suivantes sont réunies :

- Il existe au moins une instance prédite positive pour un segment étiqueté positif.
- Toutes les instances sont prédites négatives pour un segment étiqueté négatif.

Dans notre cas, la sortie du réseau de neurones est une valeur comprise entre 0 et 1. Le segment est classé positivement, si le réseau de neurones retourne une valeur supérieure ou égale à 0,5 pour au moins l'une de ses instances. En cas d'erreur, les poids du réseau sont modifiés en fonction de la sortie maximale obtenue pour toutes les instances d'un segment donné. Ainsi, nous pouvons donc, d'une part, étiqueter le segment, et, d'autre part, en cas de segment positif, retourner la position la plus vraisemblable pour le complexe K, c'est-à-dire l'instance ayant eu la valeur la plus élevée en sortie du réseau de neurones.

Au final, après l'apprentissage, il est donc possible de créer un nouvel ensemble de données, plus conforme aux problèmes traditionnels d'apprentissage. Ce nouvel ensemble d'apprentissage sera la suivant :  $\mathcal{D}_{RI} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  où chaque  $x_i$  correspond à l'instance correspondant le plus vraisemblablement à un complexe K, selon la sortie du réseau de neurones et  $y_i$  est l'étiquette associée à cette instance, déterminée en fonction de  $L_i$ , l'étiquette du segment  $B_i$ , segment duquel est extraite l'instance  $x_i$ .

### Apprentissage supervisé

Nous allons travailler ici avec l'ensemble de données  $\mathcal{D}_{RI}$  comprenant les instances représentatives sur le modèle SVM dont l'objectif est de trouver l'hyperplan qui sépare le mieux les valeurs positives et les valeurs négatives. Afin d'équilibrer les classes, et d'améliorer les performances de l'algorithme, de nouvelles instances sont créées en utilisant la méthode SMOTE [14].

Une fois l'apprentissage effectué, le réseau de neurones est utilisé sur l'ensemble de tests. Chaque instance est ainsi testée. Ainsi, pour un segment  $B'_i$ , à chaque instance  $I'_i$  est associée une étiquette  $l_i$ . Si au moins une des instances de  $B'_i$  est positive, alors le segment est étiqueté positif (il contient un complexe K). Au contraire, si toutes les valeurs sont négatives, alors le segment est étiqueté négatif (il ne contient pas de complexe K).

### Analyse du pouvoir discriminant des paramètres

Nous avons voulu tout d'abord évaluer le pouvoir discriminant de chacun des 29 paramètres proposés. À cet effet, nous avons utilisé l'algorithme *Relief* [34]. Il faut rappeler ici que les paramètres  $f_1$  à  $f_{12}$  sont des paramètres traditionnels proposés par Bankman *et al* dans [8] alors que les paramètres  $f_{13}$  à  $f_{29}$  sont nouveaux. La figure 3.36 illustre le poids affecté à chaque paramètre pour illustrer son pouvoir discriminant. Les nouveaux paramètres sont en bleu ; les standards sont en orange.

Nous pouvons observer que les nouveaux paramètres apportent une contribution importante pour la discrimination des segments positifs et négatifs. Parmi les 6 paramètres les plus discriminants, 4 sont de notre apport, principalement issus du troisième groupe permettant de comparer la partie avec des ondes « de fond » et la partie pressentie comme contenant le complexe K.

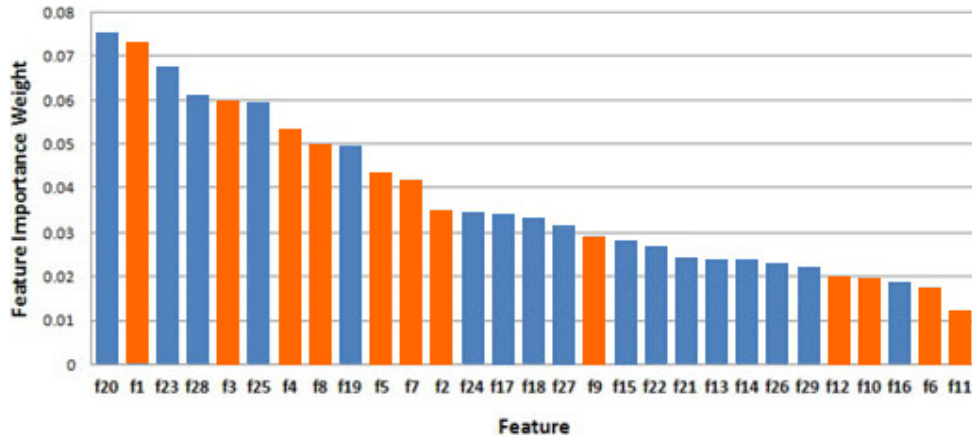


FIGURE 3.36 – Pouvoir discriminant des paramètres utilisés pour la classification

### Comparaison des 2 ensembles de paramètres

Nous avons voulu analyser l'efficacité de chacun des 2 ensembles de paramètres : le premier contenant les 14 paramètres définis dans [8] et le deuxième contenant les 29 paramètres définis précédemment. Chaque ensemble de données a été testé sur 3 algorithmes : Réseau de Neurones (ANN), Random Forest (RF) et Support Vector Machine (SVM). Les résultats de l'exactitude, la précision, le rappel et la F-Mesure sont donnés dans le tableau 3.11.

TABLEAU 3.11 – Comparaison des résultats obtenus par 3 algorithmes différents sur les 2 ensembles de données

Algorithme	Ensemble de paramètres	Exactitude	Précision	Rappel	F-Mesure
ANN	14 paramètres	0,847	0,540	0,500	0,519
	29 paramètres	0,877	0,660	0,648	0,654
RF	14 paramètres	0,853	0,588	0,370	0,455
	29 paramètres	0,874	0,676	0,463	0,549
SVM	14 paramètres	0,880	0,714	0,463	0,562
	29 paramètres	0,887	0,689	0,574	0,626

Nous pouvons constater que les exactitudes se situent au-dessus de 80% pour les 2 ensembles de paramètres sur les 3 algorithmes utilisés. Les résultats sont cependant légèrement meilleurs avec l'ensemble de 29 paramètres qu'avec celui de 14. Les algorithmes ANN et RF obtiennent une meilleure précision et un meilleur rappel sur l'ensemble de 29 paramètres. Avec l'algorithme SVM, le gain en rappel se fait aux dépens d'une perte de précision. Enfin, la F-Mesure est toujours meilleure avec l'ensemble de 29 paramètres que sur celui de 14.

### Évaluation des algorithmes

Afin d'évaluer les performances de notre algorithme, nous l'avons comparé à d'autres algorithmes. Deux comparaisons sont nécessaires car l'algorithme présenté a deux versants : le premier concerne l'apprentissage *multi-instances* et le deuxième concerne l'apprentissage *mono-instance*. La version *multi-instances* sera comparée à l'algorithme *BP-MIL* [82] ; la version *mono-instance* sera comparée à plusieurs algorithmes incluant un réseau de neurones (ANN), une machine à vecteurs de support et Random Forest (RF). Les tests sont effectués sur la plateforme Weka.

Les paramètres de ANN et RF ont été conservés par défaut, SVM a été testé avec des noyaux linéaire (*lin*) et polynomial (*pol*). RIC et BP-MIP ont été paramétrés avec un *learning rate* à 0,05.  $\delta_{pos}$  et  $\delta_{neg}$  valent respectivement 1 et 0. Le nombre d'itérations est fixé à 500. L'évaluation est la moyenne sur 10 tests successifs.

Les tests sont présentés dans le tableau 3.12.

TABLEAU 3.12 – Comparaison des algorithmes de classification

Algorithme	Exactitude	Précision	Rappel	F-Mesure
<b>BP-MIL</b>	0,874	<b>0,882</b>	0,278	0,423
<b>SVM (<i>pol</i>)</b>	0,808	0,441	0,602	0,502
<b>RF</b>	0,874	0,676	0,463	0,549
<b>SVM (<i>lin</i>)</b>	0,887	0,689	0,574	0,626
<b>ANN</b>	0,877	0,660	0,648	0,654
<b>RIC</b>	<b>0,902</b>	0,704	<b>0,704</b>	<b>0,704</b>

Nous constatons ici que BP-MIL et RF obtiennent un rappel très faible, non acceptable pour un algorithme de reconnaissance de complexes K. Malgré une exactitude, une précision et une F-Mesure correcte, pour ANN et SVM (à noyau linéaire ou non), le rappel est trop faible pour être acceptable. Avec des métriques toutes supérieures à 0,7, RIC s'avère être l'algorithme le plus efficace pour l'identification des complexes K.

### Évaluation de l'impact des valeurs $\delta_{pos}$ et $\delta_{neg}$

Afin d'évaluer l'impact de la valeur de  $\delta_{pos}$  et  $\delta_{neg}$  dans l'identification des instances représentatives avec l'algorithme BP-MIL, plusieurs valeurs ont été testées avec des valeurs centrées en 0,5 :  $\delta_{pos}$  variant de 0,5 à 1 avec un pas de 0,1 et  $\delta_{neg}$  variant de 0,5 à 0 avec un pas de 0,1. Les résultats sont représentés graphiquement sur les courbes de la figure 3.37.

La figure 3.37 montre que l'exactitude, la précision et la F-Mesure augmentent quand  $\delta_{pos}$  est proche de 1 et  $\delta_{neg}$  est proche de 0. Le rappel diminue légèrement, mais en restant compris entre 0,7 et 0,75. Il semble donc que les valeurs optimales sont  $\delta_{pos} = 1$  et  $\delta_{neg} = 0$ .

L'algorithme RIC a également été comparé à d'autres algorithmes classiques de classification de séries temporelles. En effet, les courbes EEG peuvent être assimilées à une forme de série temporelle. La reconnaissance de forme sur des séries temporelles a eu



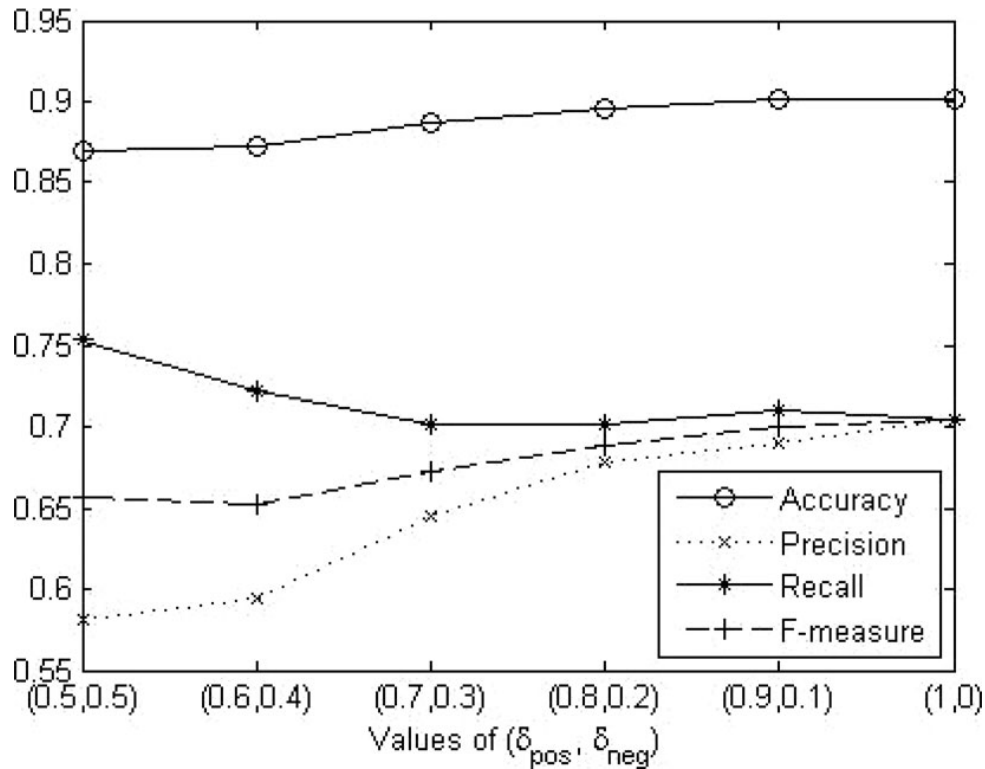


FIGURE 3.37 – Analyse de l'efficacité de l'algorithme en fonction des valeurs de  $\delta_{pos}$  et  $\delta_{neg}$

de bons résultats par des mesures de similitude, en particulier les méthodes de *déformation temporelle dynamique* (*Dynamic Time Warping, DTW*), parmi lesquelles 1NN-DTW a souvent permis d'obtenir les meilleurs résultats. Nous allons donc comparer les résultats de la classification utilisant l'algorithme RIC aux résultats de l'algorithme des 1 plus proches voisins (1NN) appliqué à des séries temporelles avec plusieurs mesures différentes : une mesure euclidienne (*Euclidean*), DTW sans contrainte, DTW avec contrainte avec une bande de *Sakoe-Chiba* à 5% et DTW avec contrainte avec une bande de *Sakoe-Chiba* à 10%. Les résultats des tests sont fournis dans le tableau 3.13.

TABLEAU 3.13 – Comparaison des algorithmes

Algorithme	Exactitude	Précision	Rappel	F-Mesure
<b>1NN-Euclidean</b>	<b>0,969</b>	<b>0,889</b>	0,148	0,254
<b>1NN-DTW (S-C 5%)</b>	0,871	0,658	0,463	0,544
<b>1NN-DTW</b>	0,887	0,718	0,519	0,603
<b>1NN-DTW (S-C 10%)</b>	0,886	0,698	0,556	0,619
<b>RIC</b>	0,902	0,704	<b>0,704</b>	<b>0,704</b>

Si ces tests confirment une amélioration des résultats en utilisant la mesure DTW par rapport à la mesure euclidienne, l'ajout de contraintes a peu d'effet sur la performance d'un algorithme de reconnaissance de complexes K sur des séries temporelles. Cependant, elle reste toujours inférieure à celle de l'algorithme RIC.



### Tests sur l'effet du suréchantillonnage

Le déséquilibre en termes de nombre d'instances des 2 classes (complexes K, non complexes K) est compensé par le recours à un suréchantillonnage utilisant l'algorithme SMOTE [14]. Nous avons donc testé plusieurs paramètres de suréchantillonnage. Nous avons donc fait varier le paramètre  $n$  qui indique le nombre de fois où la classe minoritaire est suréchantillonnée. Si  $n$  vaut 2, alors la taille de la classe est doublée, triplée pour  $n=3$ , etc. Dans nos tests, nous avons fait varier  $n$  de 1 à 10. Les résultats sont donnés sur la figure 3.38.

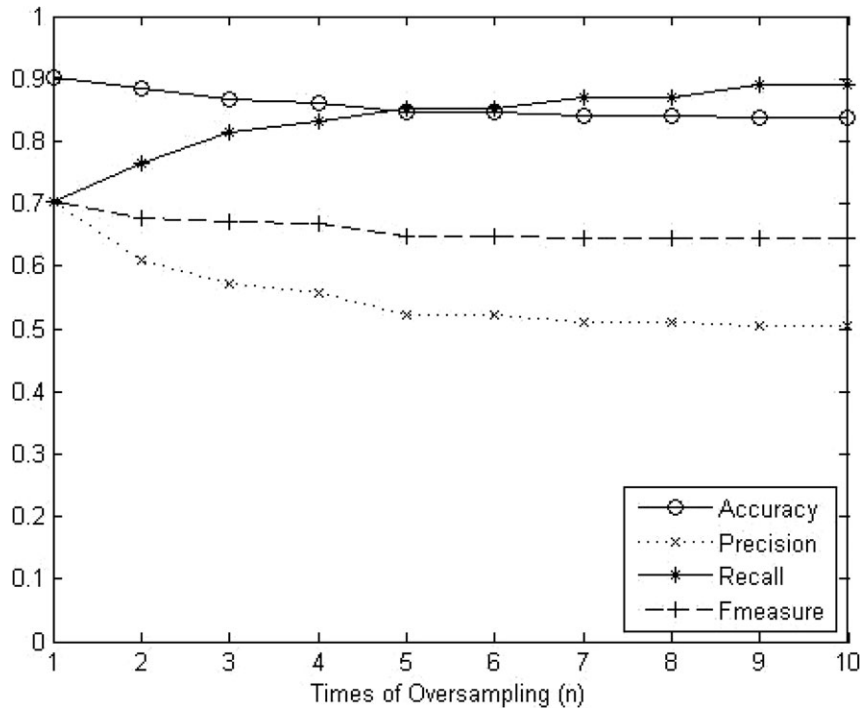


FIGURE 3.38 – Analyse de l'effet du suréchantillonnage

Nous pouvons observer sur le graphique de la figure 3.38, que la courbe de rappel augmente légèrement quand  $n$  augmente, tandis qu'au contraire, l'exactitude, la précision et la F-Mesure décroissent. De plus, les courbes se stabilisent pour  $n \geq 5$ . La courbe de rappel atteint la valeur de 0,8 dès que  $n \geq 3$ .

Ce travail a été publié dans [79].

### 3.4 Conclusion

Cette collaboration avec une unité de recherche australienne a été très riche sur le plan de l'apport mutuel de connaissances, la découverte de nouvelles techniques scientifiques, de nouvelles personnes et j'ai beaucoup apprécié l'effervescence d'idées nées de la rencontre et par l'échange de questions qui en a découlé. De nouvelles pistes de recherche ont été utilisées, en marge des approches de fusion symbolique présentées précédemment, et cela permet de constater qu'il y a des convergences dans les conclusions. En effet, une première approche a été testée pour la reconnaissance des stades de sommeil et il s'est avéré, que d'une part, la mutualisation des méthodes, pouvait améliorer les résultats, et, d'autre part,

les résultats seraient d'autant meilleurs que l'extraction de caractéristiques élémentaires serait de qualité.

## 4 Conclusion

La reconnaissance des événements respiratoires et des désaturations est une étape essentielle pour le diagnostic du Syndrome d'Apnées du Sommeil. Nous avons vu ici deux algorithmes avec cet objectif. Les désaturations sont identifiées très correctement. Les événements respiratoires sont aussi correctement détectés, mais parfois avec une étiquette différente. Il serait nécessaire de demander à un expert d'analyser les événements respiratoires scorés différemment pour qu'il aide à identifier les lacunes de l'algorithme afin de les corriger.

À l'issue de cette étape, nous disposons donc de nombreux événements singuliers qu'il faut désormais relier les uns aux autres par un lien symbolique représentant le lien de cause à effet physiologique entre leur survenue. Il s'agit également de les rapprocher d'un contexte d'occurrence pour permettre à l'expert une analyse approfondie des résultats de l'examen polysomnographique.

Premier pas vers des méthodes plus approfondies, la collaboration avec l'Université de Ballarat a été riche du point de vue de l'expérience vécue, mais aussi sur l'apport scientifique d'éléments pouvant servir à un système de support au diagnostic plus perfectionné.

# Chapitre 4

## Des caractéristiques à la décision : vers une vue globale de l'examen polysomnographique

Au début de cette étape, nous disposons du résultat du scorage (totalement automatique, semi-automatique ou totalement manuel) des données polysomnographiques, c'est-à-dire du relevé exhaustif de l'ensemble des événements physiologiques liés à la pathologie du Syndrome d'Apnées du Sommeil et à la reconnaissance des stades de sommeil. Représentés sous la forme de graphes conceptuels, ces événements sont tous isolés les uns des autres pour le moment. Nous souhaitons désormais représenter symboliquement les liens de cause à effet physiologique entre les différents événements. En effet, la survenue des événements respiratoires peut entraîner, par exemple, des événements cardiaques, des désaturations ou des micro-éveils. Nous nous intéresserons ici aux désaturations ou aux micro-éveils. N'ayant pas d'algorithme de reconnaissance automatique des micro-éveils, ils seront issus d'une analyse visuelle effectuée par l'expert médical.

### 1 Bibliographie

Il existe peu de publications sur des outils complets d'interprétation des données polysomnographiques, c'est-à-dire réalisant à la fois l'interprétation automatique des stades de sommeil, le scorage automatique des événements (respiratoires, micro-éveils et désaturations) puis la fusion de ces événements.

Néanmoins, il existe de nombreux travaux dont l'objectif est d'estimer la valeur de l'IAH. Ils consistent en général à estimer le nombre d'apnées et d'hypopnées. Le Temps de Sommeil Total utilisé est alors soit déduit d'une interprétation visuelle faite par un expert, soit estimé par la longueur de l'enregistrement avec pour hypothèse que l'acquisition des données s'est déroulée exactement sur la période de sommeil. Parfois, seules quelques heures sont analysées. L'hypothèse associée est alors que l'IAH calculé sur cette portion de sommeil est équivalent à l'IAH calculé sur l'ensemble du Temps de Sommeil.

Comme Belur V. DASARATHY l'explique dans [19], il est courant de comparer des méthodes dont la sortie se situe au même niveau de connaissance. C'est le cas ici où les travaux cités génèrent un IAH.

Les travaux peuvent être classés par source de données utilisée :

– **Électrocardiogramme et série des intervalles RR,**

En 1998, dans sa thèse de doctorat [1] et dans une publication [2], Pierre-Antoine ABSIL, de l'Université de Liège en Belgique, propose une application de méthodes d'analyse de la variabilité du rythme cardiaque au support au diagnostic du Syndrome d'Apnées du Sommeil. Cette variabilité est déduite de la série temporelle des intervalles RR calculée à partir de l'électrocardiogramme. Une première étape permet d'en extraire des motifs cycliques. Par comparaison à un seuil, ces cycles sont ensuite interprétés comme des apnées.

En 2000, Joseph MIETUS propose dans [53] une méthode qui extrait, de l'électrocardiogramme, la série temporelle des intervalles RR. Cette série temporelle est analysée en fréquence et en amplitude avec une transformée de Hilbert. La méthode utilise un apprentissage multi-instance (sur des fenêtres de 5 minutes) et vise à identifier des périodes d'apnées répétées, caractérisées par une hausse de l'amplitude et une baisse de la fréquence.

Étudiant dans l'UFR *Electrical and Computer Engineering* à l'Université d'État de Portland aux États-Unis, Wadie MARSHALL propose dans [49] une méthode de reconnaissance des apnées sur l'analyse de la variabilité du rythme cardiaque. La présence d'une apnée correspond à un pic de la variabilité du rythme cardiaque, ce qui correspond physiologiquement à une cessation puis une reprise de la ventilation.

En 2004, Ahsan H. KHANDOKER de l'Université de Melbourne en Australie a publié dans [33] une méthode de reconnaissance des apnées et des hypopnées, à partir du signal ECG découpé en fenêtres de 5 secondes, qui utilise des ondelettes. Deux étapes sont nécessaires. La première permet d'extraire les signaux pathologiques. La deuxième permet de les classer en apnées ou hypopnées.

– **Oxymétrie nocturne,**

En 1991, Jean Louis PÉPIN de l'Université de Grenoble a proposé dans [63] une méthode d'analyse de l'oxymétrie nocturne ( $SaO_2$ ) pour dépister des patients souffrant de Syndrome d'Apnées du Sommeil. Ces patients doivent ensuite passer une polysomnographie. La méthode analyse la valeur absolue de la moyenne de la dérivée du signal  $SaO_2$  sur des fenêtres de 12 secondes.

En 2005, Juan-Carlos VÁZQUEZ, de l'Université de Calgary au Canada a publié dans [78] une méthode d'estimation de l'IAH par analyse de la saturométrie nocturne en utilisant une machine à états. À chaque point enregistré inférieur au point précédent, un marqueur est noté. Lorsque la valeur de la saturation est supérieure à la précédente et qu'on dénombre au moins 3 marqueurs, alors il est considéré qu'un événement respiratoire (apnée ou hypopnée) a eu lieu.

En 2006 et 2007, Nuria OLIVER et Fernando FLOREZ-MANGAS, chercheurs chez Microsoft à Redmond aux États-Unis, proposent dans [59] et [60] un dispositif portable de diagnostic du Syndrome d'Apnées du Sommeil appelé *HealthGear*. Ce dispositif propose 2 méthodes d'estimation du nombre d'apnées. La première estime uniquement le nombre de désaturations en recalculant une nouvelle valeur de référence pour la saturation en Oxygène par fenêtres de 5 minutes et en testant plusieurs seuils de resaturation. La deuxième fait une analyse spectrale de l'oxymétrie et de la variabilité du rythme cardiaque. En cas d'apnée, un pic est observé sur le périodogramme (densité spectrale) de la saturation sur la bande de fréquences 0,015-0,04 Hz.

En 2010, J. Victor MARCOS, de l'Université de Valladolid en Espagne a présenté dans [47] une méthode à 3 niveaux d'abstraction en partant du signal issu d'une oxymétrie nocturne. La première étape extrait 2 séries de paramètres, les uns issus d'une analyse spectrale, les autres issus d'une analyse non linéaire. La deuxième étape consiste en une **Analyse en Composantes Principales** sur les 2 ensembles de données obtenus en sortie de la première étape. La troisième étape procède à une analyse discriminante linéaire sur les composantes principales issues de la deuxième étape. Le résultat est une classe qui correspond à un diagnostic.

– **autre signal unique,**

En 2005, Frédéric SENNY de l'Université de Liège en Belgique a présenté dans [68] une méthode originale de reconnaissance des événements respiratoires par une analyse des mouvements mandibulaires. Il s'agit d'une mesure indirecte des ouvertures buccales, observées chez les patients souffrant de Syndrome d'Apnées du Sommeil, souvent en fin d'apnée. Dans une première étape, sont recherchés, avec des méthodes de traitement du signal, des événements caractéristiques de la fin d'un événement respiratoire. À partir de cet événement, le temps est remonté pour déterminer une *région d'intérêt*. Cette zone est ensuite classifiée en générant deux scores, un pour l'événement intrinsèque et un autre pour le contexte de survenue. Leur somme est comparée à un seuil afin de décider si l'événement est un événement respiratoire pathologique ou non.

– **Plusieurs signaux,**

En 2001, Gabriel GUIMARÃES de l'Université Nova à Lisbonne au Portugal a publié dans [27] des travaux utilisant des techniques d'intelligence artificielle pour traiter l'aspect temporel des événements et l'appliquer à la reconnaissance des troubles respiratoires liés au sommeil. Il s'agit d'un système qui utilise plusieurs niveaux d'abstraction de l'information. Dans un premier temps, 12 paramètres sont extraits puis injectés dans des réseaux de neurones (Self Organizing Maps) afin d'obtenir des motifs primitifs. Ensuite sont identifiés des *successions* de motifs primitifs qui correspondent à une description de ce qui se passe sur la voie du flux respiratoire d'une part et sur les mouvements thoraco-abdominaux d'autre part. La survenue plus ou moins simultanée de *successions* est ensuite interprétée comme des événements. Les événements sont ensuite interprétés séquentiellement pour former des *séquences*. Ces dernières sont enfin interprétées en *motifs temporels* en fonction de leur durée. Chaque motif temporel est assimilable à un événement du domaine médical.

En 2005, Tarik AL-ANI du laboratoire Algorithmique et Architecture des Systèmes Informatiques (*A<sup>2</sup>SI*), Groupe ESIEE Paris, a publié dans [3] des travaux sur un système non invasif de classification automatique des apnées du Sommeil. Le système fait l'analyse respiratoire en utilisant 2 signaux : le flux aérien nasal et le temps de transit du pouls. Le cœur du système est un réseau de neurones dans lequel sont injectés des paramètres extraits des 2 signaux analysés.

En 2005, Paula Tamara POZZO-MENDOZA, de l'Université de Prague en République Tchèque, publie dans [66] une méthode de diagnostic du Syndrome d'Apnées du Sommeil à partir des courbes de flux aérien nasal et de temps de transit du pouls. Après un prétraitement destiné à réduire le bruit, plusieurs paramètres sont extraits de chacune des 2 courbes. La décision est prise par un réseau de neurones.

Il existe également des systèmes complets qui effectuent à la fois l'analyse respiratoire et la reconnaissance des stades de sommeil et les mettent en rapport.

Le premier système est une architecture générique de support au diagnostic médical qui utilise des observations cliniques et des résultats d'examens pour générer une proposition de diagnostic en utilisant des réseaux bayésiens. Cette méthode a été publiée dans [54]. Les résultats sont ensuite présentés à l'expert via une interface web. Cet outil a été testé sur le cas particulier du Syndrome d'Apnées du Sommeil.

Le second système s'appelle PSG-Expert. Il est présenté dans [24]. Il s'agit d'un système expert qui utilise la logique floue pour interpréter les données. Le cœur du système est un logiciel interactif de visualisation des résultats, à chacune des étapes, qui ont conduit à la décision ainsi qu'au contexte clinique du patient.

Le troisième système, s'appelle SAMOA. Présenté dans [13], il utilise des techniques d'Intelligence Artificielle. L'analyse respiratoire consiste à segmenter les cycles respiratoires, quantifier leur débit puis les classifie et les interprète. L'analyse des stades de sommeil utilise un classifieur qui prend en entrée des paramètres obtenus par traitement du signal, inspirés par les règles définies par Rechtschaffen & Kales [67]. À ces paramètres physiologiques sont ajoutés des paramètres contextuels (âge, taille, poids, ...). Au final, le système génère les indicateurs classiques :

1. l'IAH global ;
2. l'index par type d'apnées ;
3. le nombre d'événements par type ;
4. le nombre total de désaturations et resaturations ;
5. la distribution des valeurs de la saturation en oxygène ;
6. la durée de chaque stade de sommeil.

Ce système a été complété par ÁNGEL FERNÁNDEZ-LEAL dans [21] par un module de traitement de l'information temporelle. Les règles utilisées sont d'ordre quantitatif, qualitatif ou de cause à effet.

Il existe également de nombreux systèmes de support au diagnostic se comportant comme des boîtes noires, c'est-à-dire dont l'algorithme est inconnu et n'a jamais été

publié. Par exemple, le système Somnocheck, mis au point par l'entreprise allemande Weinmann, a été testé dans [73] sans que son principe soit dévoilé.

## 2 Données initiales

Les données initiales sont celles issues de l'interprétation des données polysomnographiques. Nous utiliserons l'exemple de courbes et d'événements scorés sur les courbes de la figure 4.1.

Sur cette figure, nous pouvons voir, en bleu, une cessation de flux aérien. Dans le même temps, l'effort respiratoire persiste, ce qui est caractéristique d'une apnée obstructive. Quelques secondes après la survenue de cet événement respiratoire, nous pouvons voir, d'une part, un micro-éveil, et, d'autre part, une désaturation.

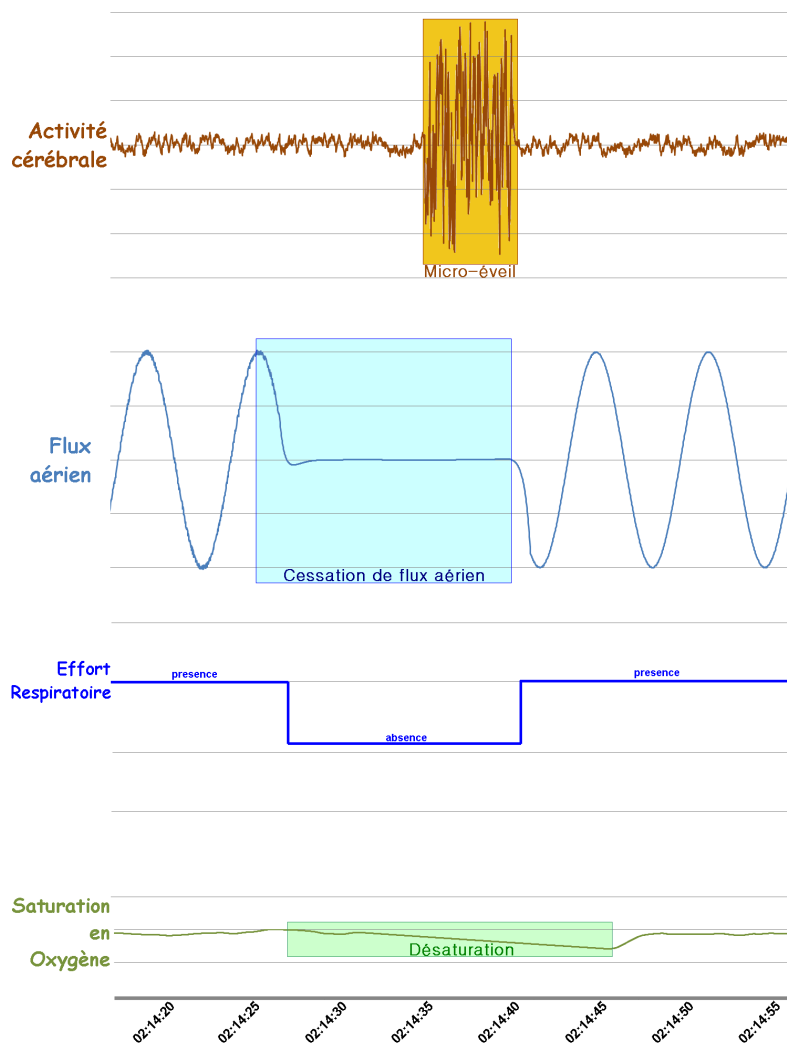


FIGURE 4.1 – Exemple de courbes et d'événements scorés

## 2.1 Données contextuelles : Stades de sommeil et Position

Le diagnostic du Syndrome d'Apnées du Sommeil analyse non seulement la fréquence d'occurrence des événements respiratoires, apnées et hypopnées, mais également leur contexte de survenue. En effet, certaines conditions, positions ou stades de sommeil, peuvent favoriser la survenue des événements respiratoires et peuvent influencer des choix de traitement.

En fonction des données enregistrées par le capteur de position, nous disposons d'un ensemble de graphes conceptuels indiquant, pour chaque période de temps sans changement de position, la position du corps (*décubitus dorsal*, *décubitus ventral*, *décubitus latéral gauche*, *décubitus latéral droit*), l'heure de début et l'heure de fin de la période. Un exemple de graphe conceptuel est donné sur la figure 4.2.



FIGURE 4.2 – Graphe conceptuel d'une période de position *décubitus dorsal*

À partir du résultat de l'interprétation des stades de sommeil, nous avons une suite d'époques avec pour chacune un stade de sommeil. Pour chaque époque, nous avons donc un graphe conceptuel. Un exemple de graphe conceptuel est donné sur la figure 4.3.



FIGURE 4.3 – Graphe conceptuel d'une époque de stade de sommeil *N2*

L'ensemble de l'enregistrement est représenté sous la forme de graphes conceptuels de manière à avoir la position et les stades de sommeil en chaque instant.

## 2.2 Données événementielles : Apnées, désaturations et micro-éveils

Nous disposons également à la suite de l'étape précédente, d'événements relevés sur les courbes polysomnographiques.

Nous utiliserons ici trois types d'événements :

- les événements respiratoires (apnées et hypopnées) ; un exemple de graphe conceptuel est donné sur la figure 4.4.
- les désaturations ; un exemple de graphe conceptuel est donné sur la figure 4.5.
- les micro-éveils ; un exemple de graphe conceptuel est donné sur la figure 4.6.



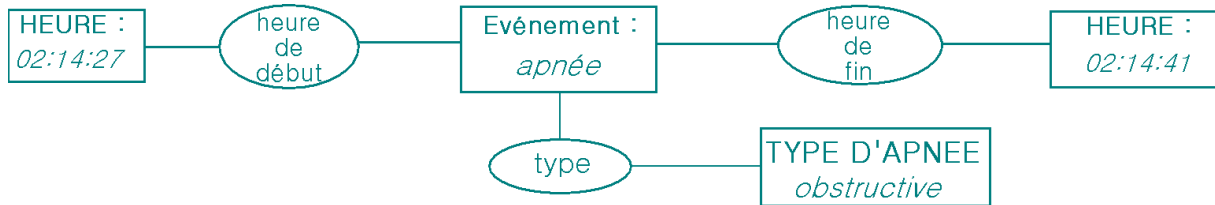


FIGURE 4.4 – Graphe conceptuel d’une apnée obstructive

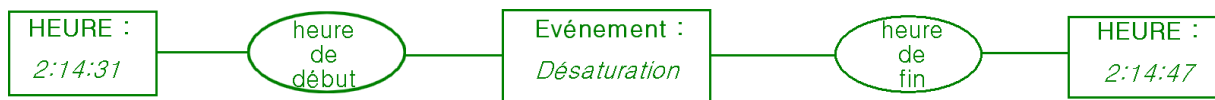


FIGURE 4.5 – Graphe conceptuel d’une désaturation

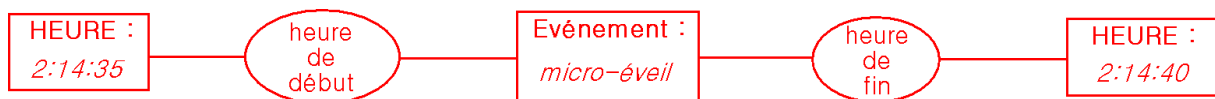


FIGURE 4.6 – Graphe conceptuel d’un micro-éveil

### 3 Méthode

La fusion s'effectue en analysant, un par un, l'ensemble des événements respiratoires. Elle consiste à rechercher, d'une part, le contexte de survenue de l'événement, et, d'autre part, ses conséquences physiologiques. Dans un premier temps, le contexte de survenue est recherché par un passage en revue des graphes conceptuels de position et de stades de sommeil. Ensuite, les conséquences physiologiques sont recherchées. Pour chaque type d'événement (non respiratoire), l'ensemble des graphes conceptuels sont analysés pour rechercher s'il existe au moins une des occurrences étant survenue au maximum 20 secondes après la fin de l'événement respiratoire. Il est alors recherché, parmi l'ensemble des graphes conceptuels d'un type d'événement donné, un événement dont la *date de début* serait située entre la *date de début* de l'événement respiratoire et 20 secondes après sa *date de fin*.

Sur l'exemple de la figure 4.1, la fusion permet d'obtenir le graphe conceptuel de la figure 4.7.

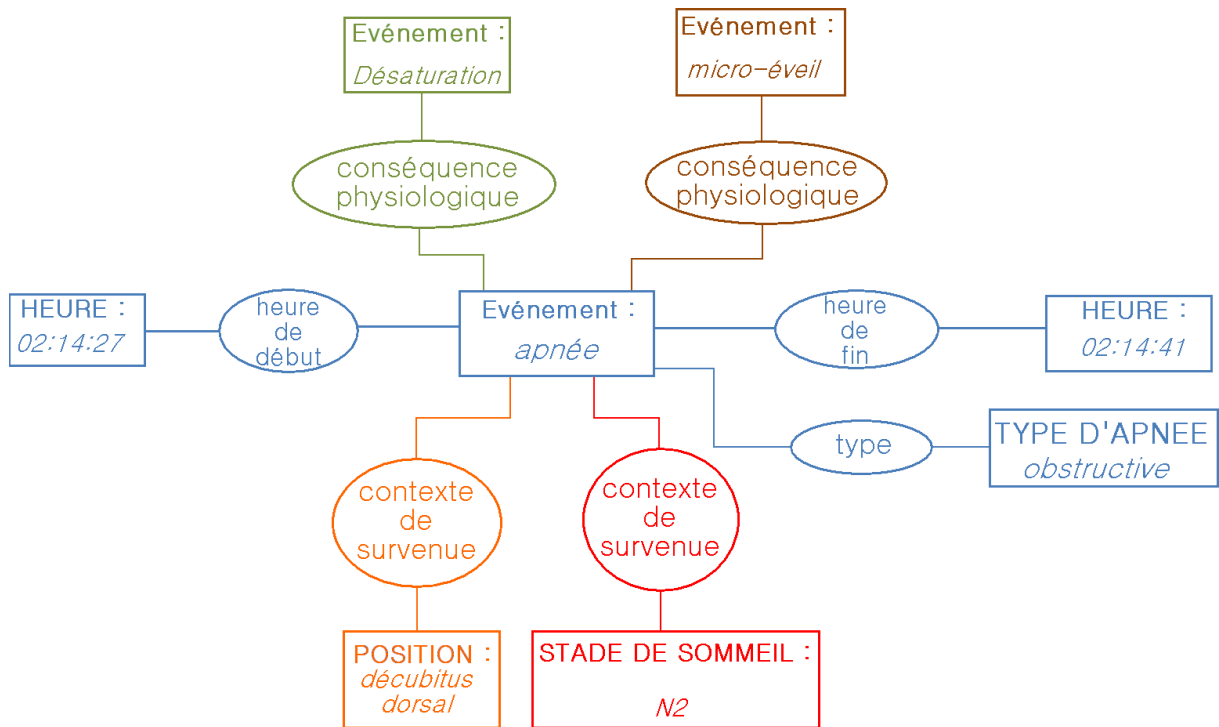


FIGURE 4.7 – Résultat de la fusion des graphes conceptuels des figures 4.2 à 4.6

### 4 Résultats

Nous avons testé cette méthode sur un ensemble de 71 dossiers d'enregistrements polysomnographiques pour lesquels nous disposons de l'ensemble des voies nécessaires sur la totalité de la durée de l'enregistrement. Les patients sont décrits dans le tableau 4.1.

Selon l'expert, l'IAH moyen de ces patients vaut  $14,54 \pm 13,9$ . Avec la méthode automatique, nous obtenons un IAH moyen de  $15,25 \pm 13,1$ . La valeur moyenne de la différence ( $IAH_{expert} - IAH_{auto}$ ) vaut  $0,72 \pm 6,84$ .

TABLEAU 4.1 – Description des patients étudiés

	Moyenne (Écart-type)	Min-Max
Âge	55 (12,5)	26-81
F/H	24/47	
Taille (m)	1,70 (0,11)	1,46-1,93
Poids (kg)	86,2 (19,5)	46-177
IMC ( $kg/m^2$ )	29,9 (6,0)	16,3-51,2
IAH	14,5 (13,9)	0-58

La figure 4.8 illustre graphiquement l'IAH trouvé par l'expert (en abscisse) et l'IAH obtenu par la méthode automatique (en ordonnée). L'interprétation est facilitée par la présence de l'axe médian et l'affichage de la classe de sévérité de chaque patient, symbolisée par un marqueur spécifique; les zones du graphique correspondant à ces classes de sévérité sont mises en évidence par un fond coloré.

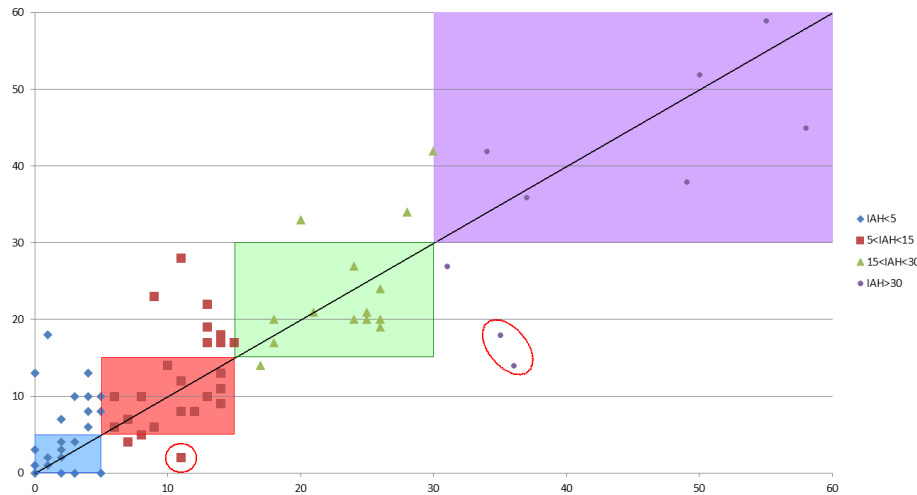


FIGURE 4.8 – Comparaison des IAH de l'expert et obtenus par l'analyse automatique pour l'ensemble des patients

Nous pouvons y remarquer que la plupart des points sont proches de l'axe médian, ce qui indique que la méthode utilisant l'analyse automatique obtient des résultats très proches de ceux obtenus par l'expert par une analyse visuelle des tracés.

Globalement, la méthode permet de retrouver la même classe de sévérité que l'expert, ou la classe de sévérité supérieure. Il existe 3 patients pour lesquels la classe de sévérité est sous-estimée, tout en restant très proche de la frontière de la classe obtenue par l'expert. Seuls trois points, mis en évidence par des cercles rouges, sont nettement sous-estimés par rapport à l'expert. Nous nous proposons d'étudier ces 3 cas plus en détails.

Le premier cas est celui du patient symbolisé par un carré rouge, symbolisant un SAS léger. Il s'agit du patient 2948. D'après l'expert, ce patient a un IAH de 11; selon l'analyse automatique, l'IAH vaut 4. L'expert avait identifié 17 apnées. Parmi elles, 8 ont été correctement scorées, 5 ont été identifiées comme des hypopnées et 4 n'ont pas été détectées. Ces 4 apnées non détectées sont représentées graphiquement sur la figure 4.9. L'expert a également scoré 109 hypopnées, parmi lesquelles 96 ont été également correctement détectées et identifiées par l'analyse automatique, les 12 autres n'ayant pas

été détectées. Ces 12 hypopnées sont représentées graphiquement sur la figure 4.10.

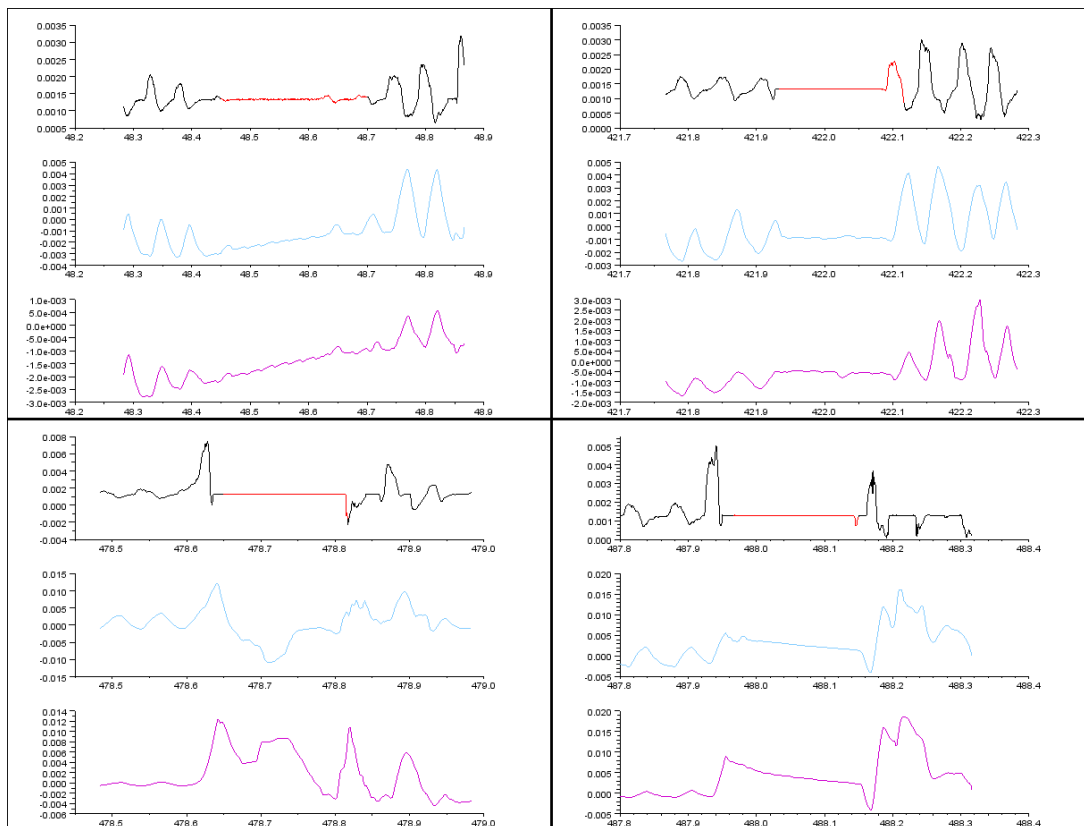


FIGURE 4.9 – Apnées du patient 2948 non détectées par l'analyse automatique

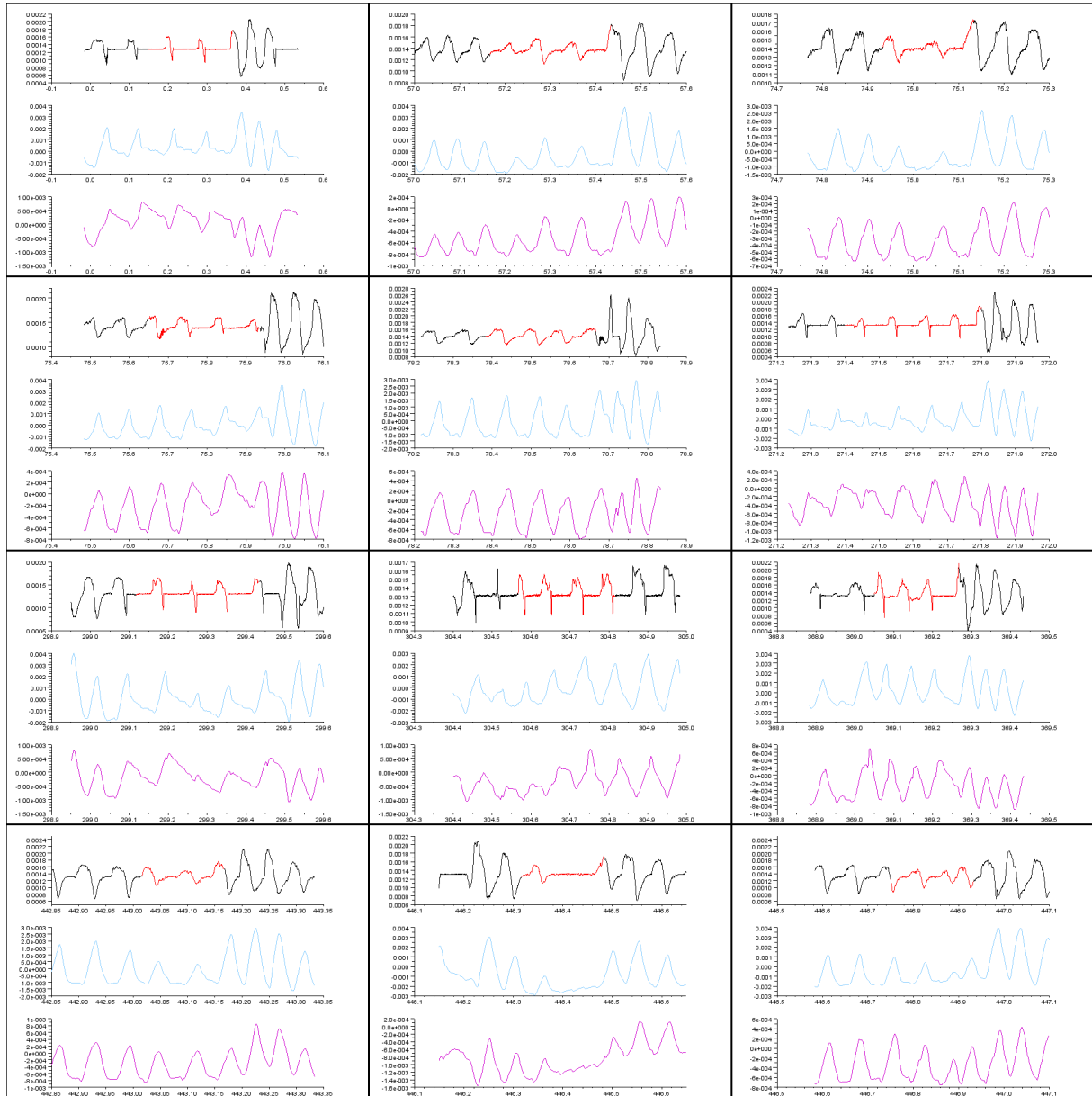


FIGURE 4.10 – Hypopnées du patient 2948 non détectées par l'analyse automatique

Le deuxième cas est celui du patient 64. D'après l'expert, ce patient a un IAH de 35 ; d'après l'analyse automatique, l'IAH vaut 18. L'expert avait identifié 56 apnées. Parmi elles, 12 ont été également scorées comme des apnées, 43 ont été scorées comme des hypopnées et 1 n'a pas été détectée par l'analyse automatique. Elle est illustrée sur la figure 4.11. Sa durée selon les critères de l'AASM [31] est estimée à 10 secondes et 20 centièmes. Comme le montre l'extrait du rapport d'événements scorés par l'expert au format XML fourni en figure 4.12, une hypopnée a été scorée sur une période qui recouvre celle de l'apnée non détectée par l'analyse automatique. Il semble donc que le scoreur humain ait eu recours à une préanalyse et ait, en conséquence, surestimé l'IAH. De plus, l'expert avait identifié également 174 hypopnées. Parmi elles, 119 ont correctement été identifiées ; les 55 autres n'ont pas été détectées par l'analyse automatique. Une sélection de 6 de ces hypopnées est donnée sur la figure 4.13.

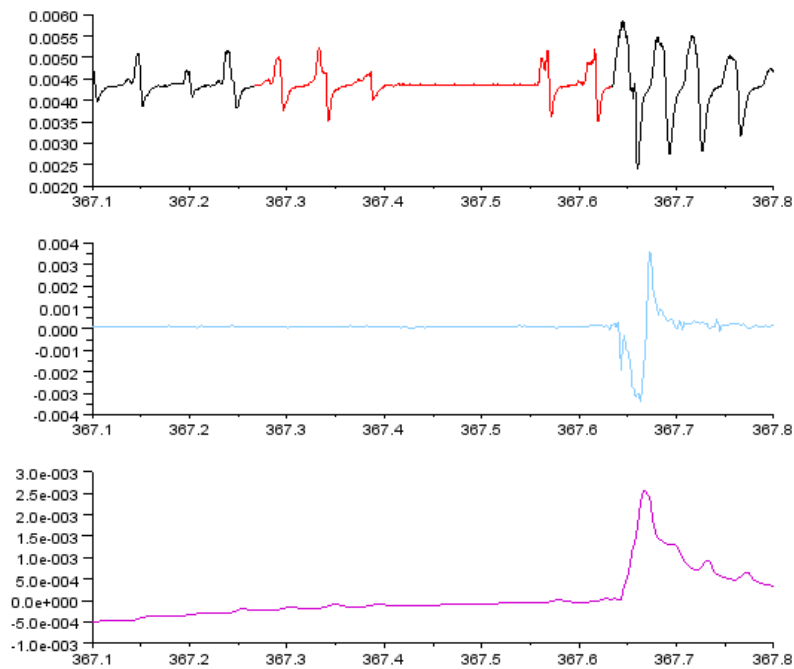


FIGURE 4.11 – Apnée du patient 64 non détectée par l'analyse automatique

```

- <Event>
  <Type dt:dt="string">APNEA-MIXED</Type>
  <Location dt:dt="string">Resp.Pressure-Cannula.Nasal</Location>
  <StartTime dt:dt="string">2007-09-26T06:07:16.337283</StartTime>
  <StopTime dt:dt="string">2007-09-26T06:07:37.866983</StopTime>
</Event>
- <Event>
  <Type dt:dt="string">DESAT</Type>
  <Location dt:dt="string">SpO2.Averaged-Probe</Location>
  <StartTime dt:dt="string">2007-09-26T06:07:19.981210</StartTime>
  <StopTime dt:dt="string">2007-09-26T06:07:49.981210</StopTime>
  + <Parameters></Parameters>
</Event>
- <Event>
  <Type dt:dt="string">HYPOPNEA</Type>
  <Location dt:dt="string">Resp.Flow-Cannula.Nasal</Location>
  <StartTime dt:dt="string">2007-09-26T06:07:22.531210</StartTime>
  <StopTime dt:dt="string">2007-09-26T06:07:34.631210</StopTime>
</Event>

```

FIGURE 4.12 – Extrait du rapport d'événements du patient 64 scorés par l'expert (au format XML)

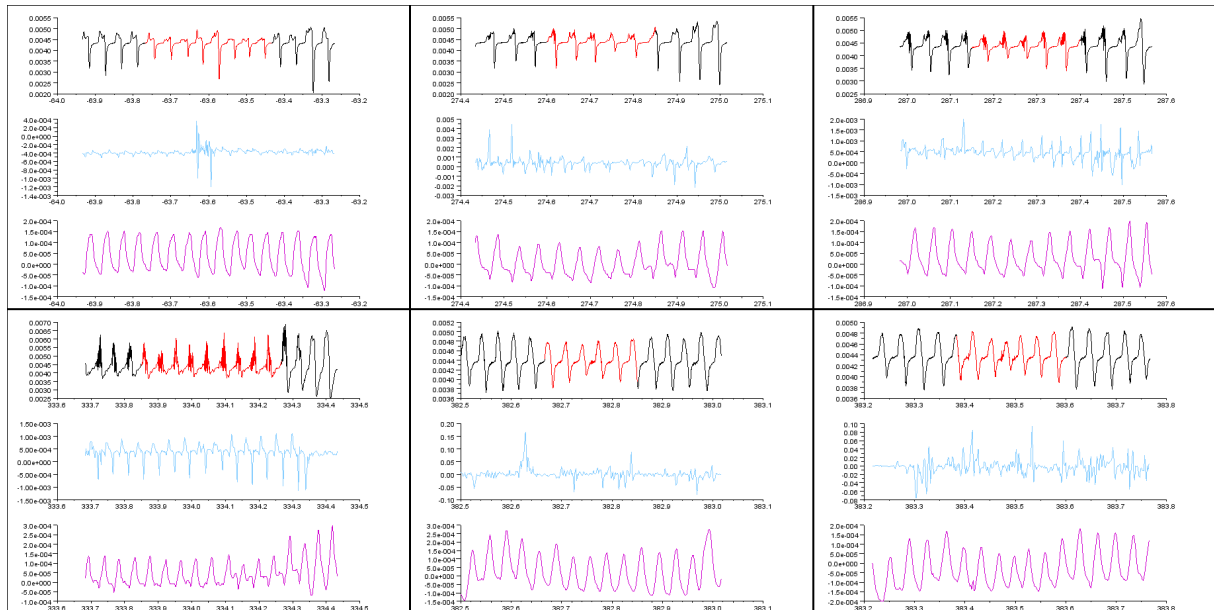


FIGURE 4.13 – Hypopnées du patient 64 non détectées par l'analyse automatique

Le troisième cas est celui du patient 1292. Pour ce patient, l'expert a obtenu un IAH de 36 alors que l'analyse automatique conclut à un IAH de 14. Cette différence, la plus importante des 3 cas étudiés, s'explique notamment par le fait que le signal est très bruité. Nous pouvons le constater sur la figure 4.14 qui représente le signal enregistré de pression nasale sur 3 heures avec un zoom sur 3 minutes.

L'expert avait détecté 141 apnées et 55 hypopnées, parmi lesquelles 52 apnées et 34 hypopnées n'ont pas été détectées par l'analyse automatique, principalement à cause du caractère particulièrement bruité du signal permettant de mesurer le flux aérien nasal.

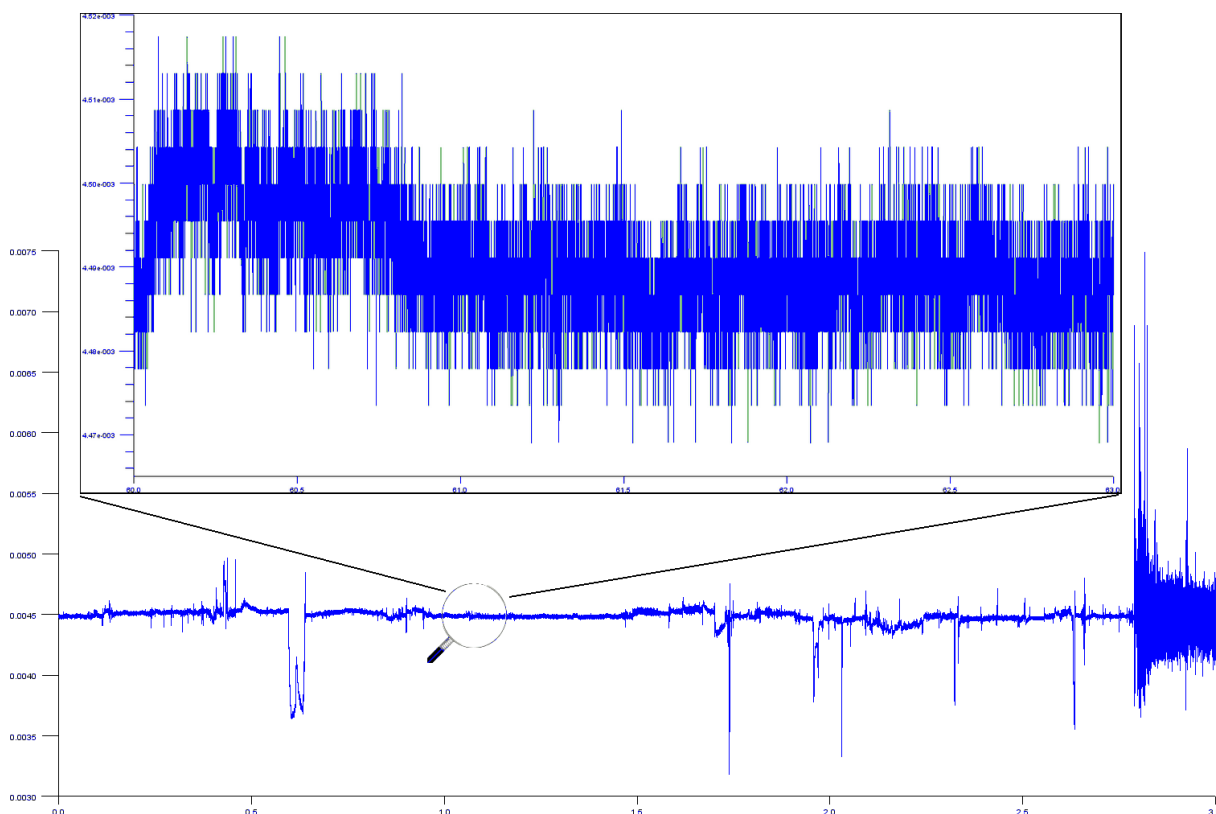


FIGURE 4.14 – Signal de pression nasale du patient 1292

Pour étudier les valeurs de l'IAH automatique, nous nous proposons de construire un diagramme en *boîte à moustaches* (ou *boxplot*) pour chaque classe de sévérité (déduite de l'IAH de l'expert). Ce graphique est représenté sur la figure 4.15. Pour faciliter son interprétation, les zones de valeur de l'IAH correspondant à chaque classe de sévérité sont représentées sur fond jaune.

Nous pouvons y remarquer que pour les classes de sévérité faible ( $IAH \leq 5$  et  $5 < IAH \leq 15$ ), l'analyse automatique donne des résultats similaires ou une sévérité légèrement supérieure. Pour la classe des SAS modérés ( $15 < IAH \leq 30$ ), contenant 14 patients, l'analyse automatique a également globalement trouvé la même classe que l'expert, avec uniquement quelques cas surestimés. En ce qui concerne la classe de SAS sévères ( $IAH \geq 30$ ), contenant 9 patients, l'analyse automatique aboutit également à la même classe de sévérité. Nous remarquons cependant que pour 1/3 des patients, la sévérité est sous-estimée. Il s'agit de 3 patients : les patients 64 et 1292, étudiés en détails précédemment, et un troisième patient, pour lequel l'expert a trouvé un IAH de 31,



tandis que l'analyse automatique a trouvé un IAH de 27. Il s'agit donc bien de classes de sévérité différentes, mais proches l'une de l'autre, avec des valeurs d'IAH situées de part et d'autre du seuil entre les classes de SAS modérés et SAS sévère.

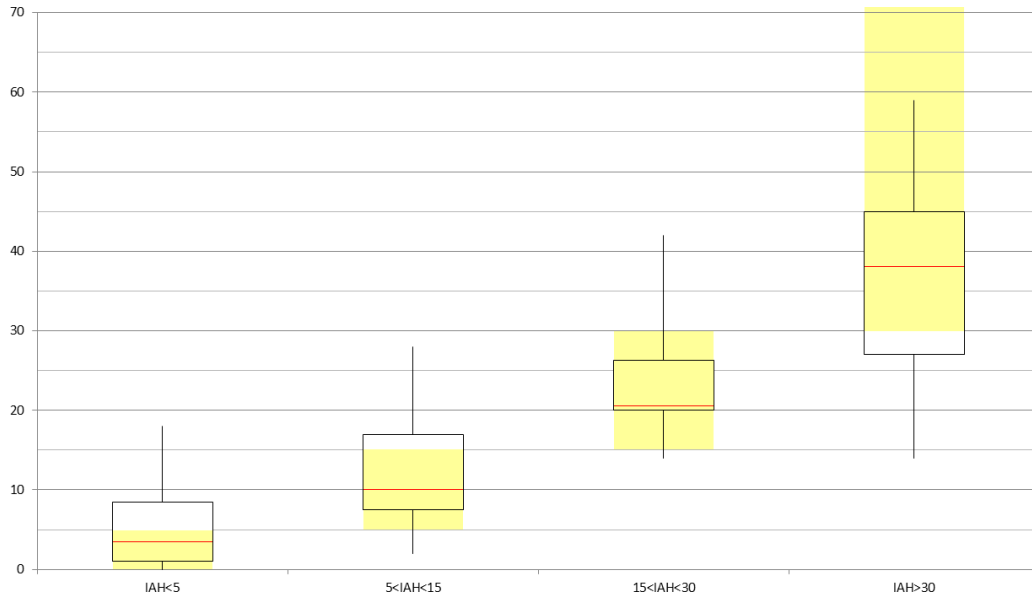


FIGURE 4.15 – Boîtes à moustaches, par classe de sévérité, des IAH obtenus par l'analyse automatique

Il apparaît qu'il est nécessaire de mettre en regard l'IAH de l'expert à celui obtenu par l'analyse automatique. Le raisonnement par classe de sévérité peut être trompeur. Deux IAH proches, mais situés de part et d'autre de la frontière entre deux classes aggrave l'erreur réelle. De même, deux IAH situés dans la même classe de sévérité, mais proches des deux extrémités de l'intervalle de valeurs de la classe seraient considérés comme valides, et éventuellement cacher des écarts importants.

Nous nous proposons donc d'étudier les boîte à moustaches, par classe de sévérité du pourcentage d'écart entre les deux IAH calculé selon la formule suivante :

$$\frac{|IAH_{auto} - IAH_{expert}|}{IAH_{expert}}$$

Ce graphique est représenté sur la figure 4.16. Ce premier graphique met en évidence des écarts très importants pour la classe la moins sévère, notamment pour un patient, dont le pourcentage d'écart entre les IAH est de 1700%. Ce patient, portant l'ID 1883, a un IAH de 1 selon l'expert et de 17 selon l'analyse automatique. Cela signifie que l'expert a conclu que ce patient n'avait pas de SAS significatif alors que l'analyse automatique conclut à un SAS modéré. Pour faire une analyse plus précise, il convient de zoomer sur le bas du graphique, ce qui est représenté sur la figure 4.17.

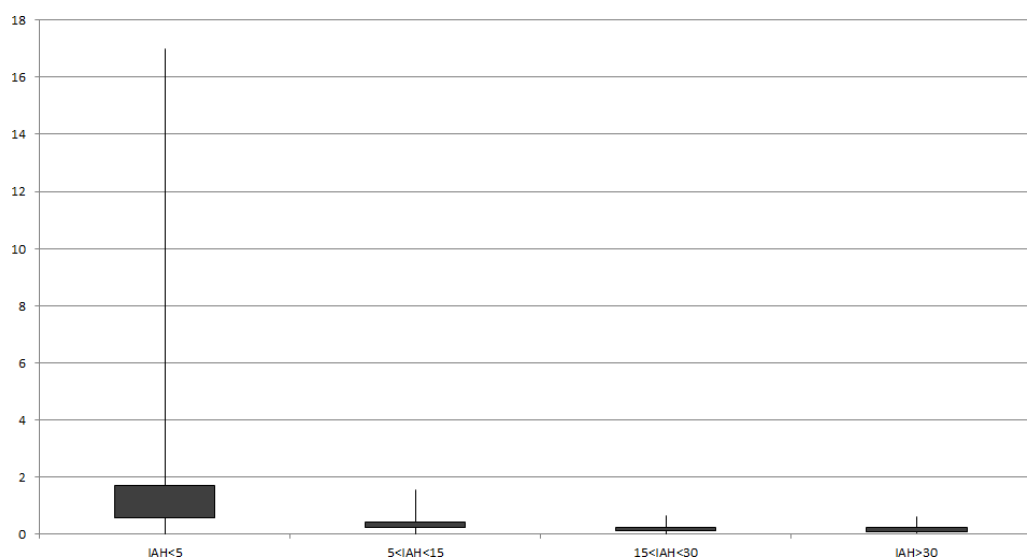


FIGURE 4.16 – Boîtes à moustaches, par classe de sévérité, du pourcentage d'écart entre les IAH (graphique complet)

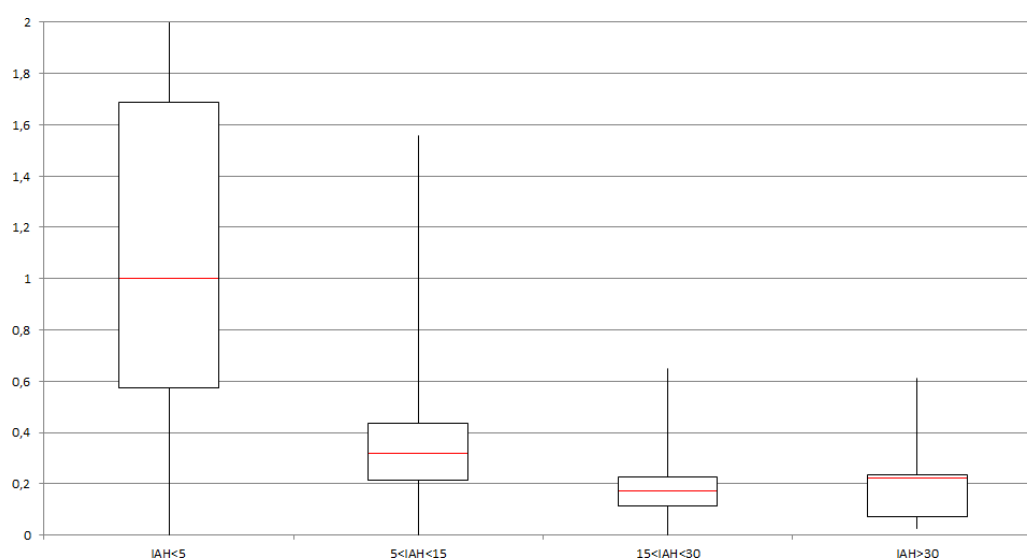


FIGURE 4.17 – Boîtes à moustaches, par classe de sévérité, du pourcentage d'écart entre les IAH (zoom sur la partie inférieure)

Nous observons que pour la classe des SAS non significatifs, l'écart entre les IAH se situe entre 60% et 160%, c'est-à-dire autour de 100%. Cela peut paraître grand, mais compte tenu des faibles valeurs d'IAH pour cette classe ( $IAH < 5$ ), le pourcentage d'erreur est plus sensible, mais reste acceptable, tant qu'il ne conduit pas à une conclusion différente, ce qui est le cas, comme nous avons pu le voir précédemment.

Pour les autres classes de sévérité, le pourcentage d'écart se situe entre 15% et 40%, en dehors des cas particuliers étudiés à part. Ceci reste acceptable pour un outil de support au diagnostic. Pour analyser ce résultat, il faut le comparer aux différences entre les IAH obtenus par des scoreurs humains. Dans [72], Masaaki SUZUKI mentionne une étude comparative des scorages d'un même dossier de courbes polysomnographiques par 16 techniciens issus de laboratoires du sommeil différents au Japon, tous expérimentés dans l'interprétation de signaux polysomnographiques. Au final, les valeurs de l'IAH variaient de 13,5/h à 29,5/h. Le coefficient de variation a également été mesuré. Il est calculé par le rapport entre l'écart-type et la moyenne sur les 16 valeurs de l'IAH. Dans cette étude, le coefficient de variation valait 24,8%. En conclusion, le résultat de l'analyse automatique semble donc se situer dans le même ordre de grandeur que celui obtenu par plusieurs techniciens humains, et valide ainsi la méthode utilisée.

Intéressons-nous maintenant au cas du patient 1883, pour lequel l'expert a conclu à un IAH de 1 alors que l'analyse automatique conclut à un IAH de 18. Chez ce patient, l'analyse automatique a détecté un total de 40 apnées et 200 hypopnées. Parmi elles, 32 apnées et 5 hypopnées étaient en dehors de la période d'analyse de l'expert ; 8 apnées et 187 hypopnées n'ont pas été relevées par l'expert ; les 8 autres hypopnées avaient également été scorées par l'expert. La première remarque concerne la qualité du signal. En effet, comme le montre la figure 4.18, le signal est très perturbé. La période d'analyse y est représentée en rouge, commence à 22h23 et se termine à 6h46mn30s. En dehors de cette période, le signal n'est pas interprétable, ce qui invalide les 32 apnées et 5 hypopnées détectées par l'analyse automatique en dehors de la période d'analyse de l'expert. Les 8 apnées non relevées sont représentées sur la figure 4.19. Il apparaît que la plupart de ces apnées ne répondent pas aux critères de l'AASM [31], l'erreur provenant de la mauvaise qualité du signal. Nous avons choisi 9 hypopnées détectées par l'analyse automatique, non relevées par l'expert et les avons représentées sur la figure 4.20. Ces hypopnées semblent répondre aux critères de l'AASM. Il faudrait donc les soumettre à un expert pour éventuellement les valider. Ceci n'a pas été fait au cours de ce doctorat.

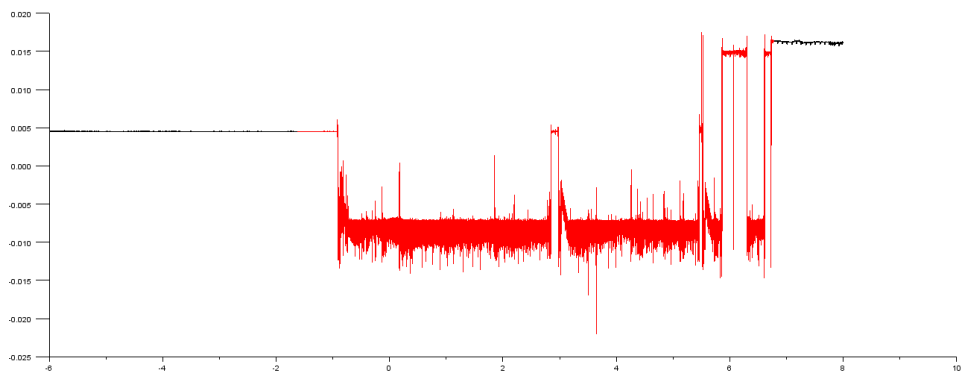


FIGURE 4.18 – Flux respiratoire du patient 1883

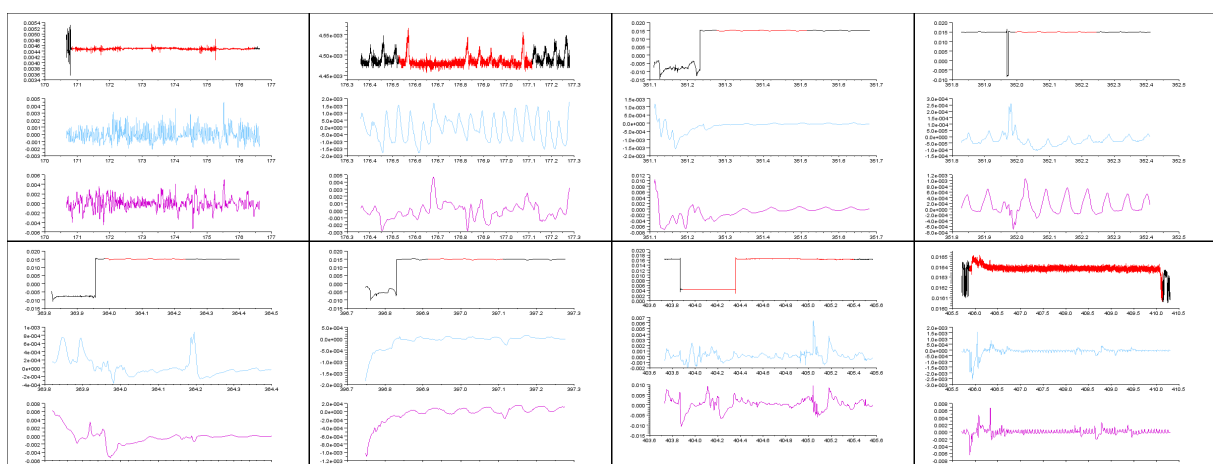


FIGURE 4.19 – Apnées non détectées du patient 1883

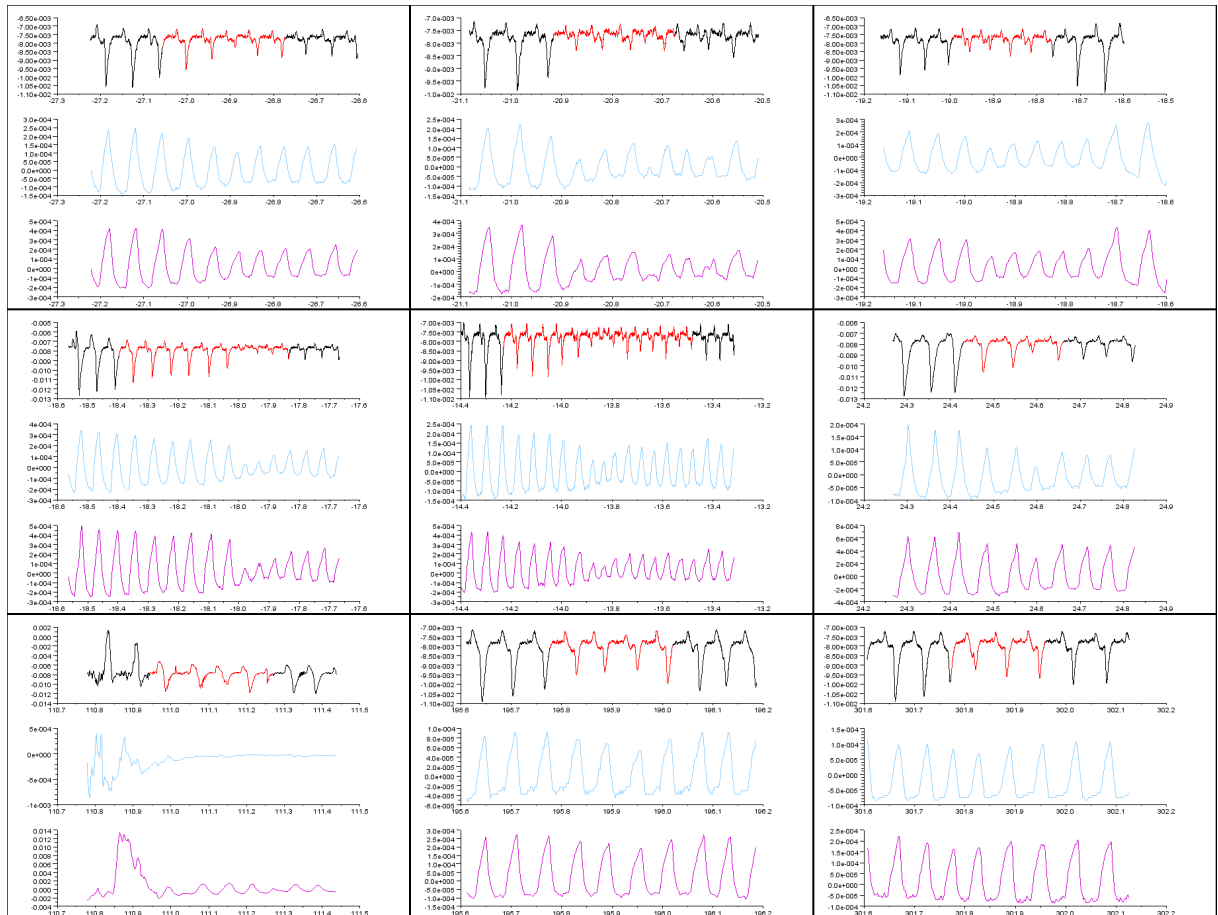


FIGURE 4.20 – Sélection de 9 hypnognées non détectées chez le patient 1883

## 5 Discussion

Dans l'analyse de notre méthode, nous sommes confrontés à une première limite : il est difficile d'évaluer le résultat final de la fusion, en incluant la validité et l'intérêt de la mise en correspondance des causes et des effets physiologiques parmi les événements détectés ainsi que leur double contexte de survenue. Un compte-rendu polysomnographique classique inclut au mieux les index de survenue des apnées et des hypopnées par position puis par stade de sommeil, mais jamais les deux à la fois. Il s'agit d'un apport tout à fait nouveau pour un médecin spécialiste des troubles du sommeil. Il faudrait donc en toute rigueur procéder à une évaluation comparant les pratiques (décisions médicales et traitements proposés) des praticiens, et, éventuellement aussi l'opinion du patient sur sa prise en charge en utilisant 2 groupes de patients : le premier groupe de patients suivant la prise en charge habituelle (de référence), le second serait pris en charge exclusivement avec notre méthode, le résultat étant soumis au praticien pour la prise de décision sur le diagnostic et la proposition de traitement. Il faudrait alors trouver des indicateurs de comparaison ; l'observance pourrait peut-être en faire partie.

Pour évaluer correctement l'interprétation seule des courbes, nous avons vu qu'il était imparfait de comparer l'IAH global, ou même les événements un à un, soit par l'expert médical, soit par l'analyse automatique. En effet, un autre expert obtiendrait une interprétation différente. Une évaluation rigoureuse pourrait consister à pré-analyser les données avec notre analyse automatique et à soumettre le résultat à deux experts différents pour correction. L'évaluation finale consisterait alors à mesurer le taux d'accord entre les 2 experts après leurs corrections respectives, mais aussi le taux de corrections apportées par rapport au nombre d'événements relevés chez ce patient.

Le traitement automatique, partiel ou intégral, des données polysomnographiques est un problème complexe. Nous avons vu que l'approche qui est généralement choisie est de réduire la complexité en n'utilisant que quelques courbes, parfois une seule. Il existe aussi des systèmes complets, mais la plupart d'entre eux est protégée et leur fonctionnement est inconnu. De plus, les experts différencient clairement les événements respiratoires, les désaturations et les événements cardiaques comme l'indique la liste des indicateurs à inclure dans un compte-rendu polysomnographique cités par l'AASM [31], alors que beaucoup d'approches estiment l'IAH par une mesure de l'index de désaturations ou d'événements cardiaques.

L'un des intérêts de notre approche est de proposer une analyse complète qui utilise plusieurs modules, chacun pouvant être soumis à une évaluation et ensuite être amélioré. Il pourra donc être envisagé de compléter l'analyse des courbes par des données cliniques. L'âge, la taille, le poids (et donc l'Indice de Masse Corporelle ou IMC), ou le sexe sont des facteurs susceptibles de jouer sur le diagnostic et/ou le traitement. La fusion symbolique est une méthode efficace quand les sources sont hétérogènes. D'autres modules, de reconnaissance automatique des événements cardiaques, en particulier la variabilité du rythme cardiaque, ou des mouvements de jambes ainsi que le calcul automatique de la pléthysmographie de pouls sont des modules qui pourraient apporter une contribution à rendre l'outil plus fiable et plus précis.

Nous avons montré ici que notre approche permet une estimation de l'IAH tout à fait comparable à celle qui pourrait être obtenue par un autre expert. La classe de sévérité

est correctement respectée. Afin de prévenir des erreurs, un prétraitement serait nécessaire pour évaluer la quantité de bruit présent sur le signal et éventuellement le réduire.

Il convient cependant de comparer nos résultats à ceux obtenus par les autres méthodes. Pour être au plus juste du résultat attendu par le médecin, nous allons nous concentrer sur les travaux qui estiment l'IAH, ce qui est finalement assez rare. En effet, beaucoup d'études citées précédemment cherchent à dépister les personnes malades ou non et affichent la sensibilité et la spécificité. Pour cela, ils doivent au préalable décider d'un seuil d'IAH au-dessus duquel ils considèrent que le SAS est pathologique.

Dans [78], Juan-Carlos VÁZQUEZ de l'Université de Calgary au Canada a testé son approche sur 245 patients ( $IAH = 25,6 \pm 16,8$ ). Avec sa méthode, en moyenne,  $IAH_{auto} - IAH_{expert}$  vaut 2,18 et l'écart-type 6,17.

Dans [33], Ahsan H. KHANDOKER teste sa méthode sur 16 patients ( $IAH = 16,06 \pm 19,5$ ). En moyenne, son algorithme obtient un IAH de 14,86 avec un écart-type de 16,9. La valeur moyenne de la différence des IAH entre l'analyse automatique et l'expert ( $IAH_{auto} - IAH_{expert}$ ) vaut -1,20; l'écart-type vaut 3,9.

En comparaison, nous avons 71 patients ( $IAH = 14,54 \pm 13,9$ ). L'analyse automatique donne le résultat suivant :  $IAH = 15,25 \pm 13,1$ . La moyenne de ( $IAH_{auto} - IAH_{expert}$ ) vaut 0,72 avec un écart-type de 6,84.

En retirant les résultats du patient 64, dont le rapport d'événements montre des recouvrements, et ceux du patient 1292 dont le signal est très bruité, la moyenne de ( $IAH_{auto} - IAH_{expert}$ ) vaut alors 1,3 et l'écart-type vaut 5,97.

Nous pouvons voir que l'IAH est mieux estimé par notre méthode que par celle de J.-C. VÁZQUEZ. Les résultats sont comparables à ceux de Ahsan H. KHANDOKER mais en surestimant légèrement l'IAH alors que Ahsan H. KHANDOKER le sous-estime. Ceci répond davantage aux attentes d'un expert car il est plus facile d'invalider des événements scorés de façon non pertinente, plutôt que d'en supprimer. Cette fiabilité est renforcée par le fait que notre méthode effectue la reconnaissance automatique des stades de sommeil, ce qui n'est pas le cas des 2 autres. Elle calcule donc l'IAH avec une meilleure estimation de la durée de sommeil.

Dans [73], Ana Claudia TONELLI DE OLIVEIRA teste le dispositif SomnoCheck sur 149 patients. Elle calcule la sensibilité et la spécificité de la méthode en utilisant plusieurs seuils différents pour considérer le SAS comme pathologique. Les résultats sont donnés dans le tableau 4.2 et comparés aux nôtres.

TABLEAU 4.2 – Sensibilité et la Spécificité en utilisant la méthode de A. C. TONELLI DE OLIVEIRA avec plusieurs seuils d'IAH de SAS pathologique

seuils d'IAH	TONELLI DE OLIVEIRA		UGON (71 patients)		UGON (69 patients)	
	Sensibilité	Spécificité	Sensibilité	Spécificité	Sensibilité	Spécificité
$IAH \geq 5$	96,15	64,7	94,0	61,9	93,8	61,9
$IAH \geq 15$	81,3	82,6	91,7	83,0	95,5	83,0
$IAH \geq 30$	80,0	92,1	70,0	96,7	87,5	96,7

L'expert médical est intéressé en particulier pour les cas modérés ( $15 \leq IAH < 30$ ) et encore plus par les cas sévères ( $IAH \geq 30$ ). En retirant les 2 patients 64 et 1292, notre méthode obtient une sensibilité et une spécificité supérieures et s'avère donc de meilleure qualité que le dispositif SomnoCheck.

## 6 Conclusion

Nous avons démontré ici que la fusion symbolique est efficace pour traiter les problèmes biomédicaux, en particulier celui du support au diagnostic du Syndrome d'Apnées du Sommeil à partir des courbes polysomnographiques. L'utilisation du formalisme des graphes conceptuels permet de travailler indépendamment des méthodes de recueil des événements et de scoring automatique des stades de sommeil. Un autre avantage de la fusion symbolique est qu'il est possible de la compléter par d'autres sources d'informations, éventuellement hétérogènes. Ainsi, l'utilisation d'un algorithme de reconnaissance automatique des micro-éveils, des mouvements de jambes ou des événements cardiaques permettrait d'enrichir le résultat en utilisant sans changer la méthode.

Sur la problématique spécifique du Syndrome d'Apnées du Sommeil, les tests ont permis de montrer que les résultats sont au moins équivalents ou meilleurs que ceux des méthodes déjà existantes. Nous avons démontré qu'ils sont comparables à ceux obtenus par un expert humain. L'analyse des erreurs permet de voir que la qualité de la fusion est dépendante de la qualité des sources fusionnées. Les écarts les plus importants sur l'IAH calculé par l'analyse automatique proviennent généralement soit d'un relevé imparfait des événements à fusionner, ceci à cause du bruit présent sur les courbes enregistrées, soit par la mise en évidence d'un IAH de l'expert discutable, sans doute dû à l'utilisation d'une pré-analyse automatique des signaux.

L'évaluation de la fusion symbolique reste partielle car il n'existe pas aujourd'hui de référence de scoring liant les événements entre eux comme il est effectué dans notre approche. Ceci pose la question de leur utilisabilité. Il devient nécessaire de construire une visualisation apte à présenter à l'expert le contenu sémantique inhérent aux indicateurs diffusés. C'est ce que nous nous proposons de faire dans le prochain chapitre.



## Chapitre 5

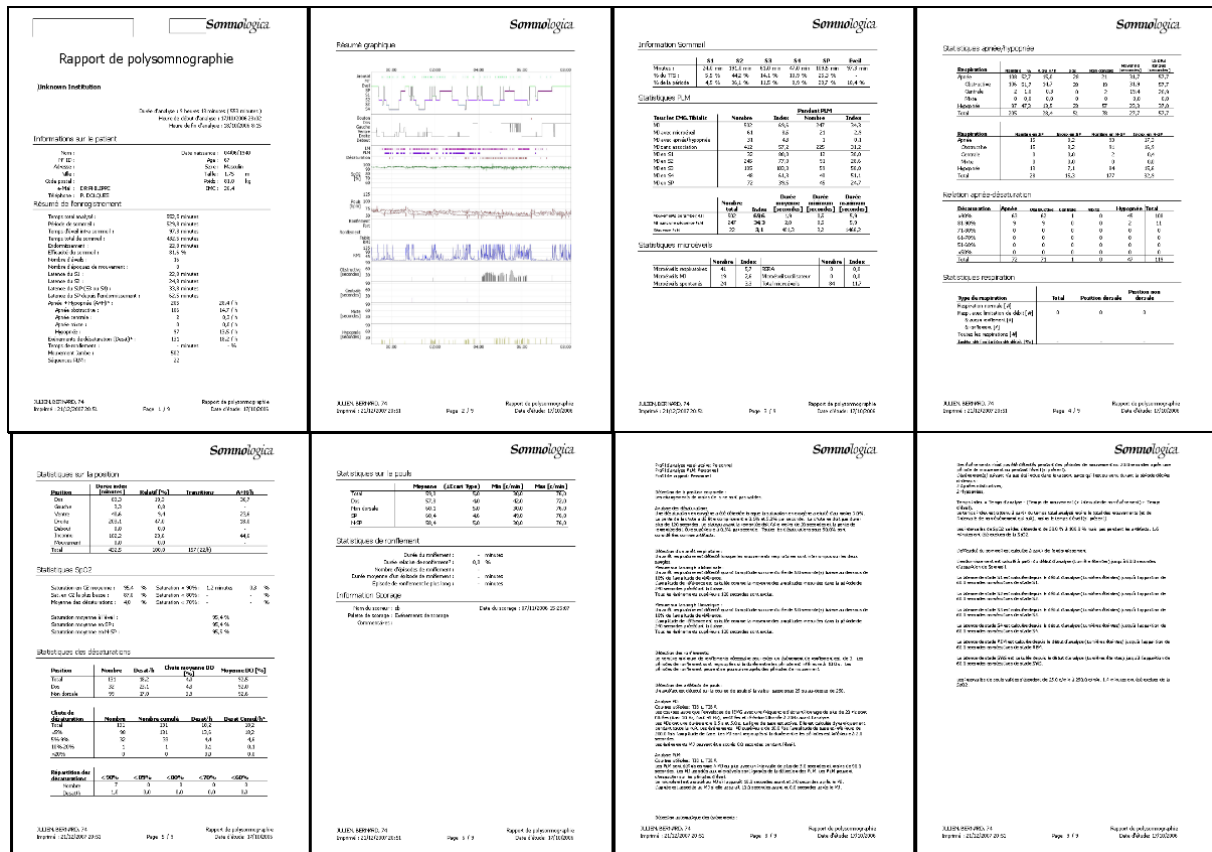
# Extraction du contenu sémantique par visualisation

Une des problématiques que nous avons une fois les résultats de la fusion symbolique obtenus, c'est de les présenter au médecin avec une représentation suffisamment riche sémantiquement parlant pour pouvoir faire un diagnostic précis et rapide. Dans notre cas, la méthode génère 256 index différents pour les apnées et 16 pour les hypopnées. Il est impossible d'intégrer leur contenu sémantique sans utiliser une méthode de visualisation.

Il s'y ajoute le problème de la communication des résultats. En effet, la médecine du sommeil est un domaine médical où « cohabitent » des experts de spécialités différentes (neurologie et pneumologie du sommeil) qui travaillent régulièrement en collaboration avec d'autres experts (médecins généralistes, spécialistes en orthopédie maxillo-faciale, cardiologie, neurologie, pneumologie ou en psychiatrie). Il est important que ces experts se comprennent le mieux possible et aient la même interprétation des mêmes résultats. Nous allons donc chercher ici à trouver une présentation des résultats qui allie simplicité, richesse et homogénéité d'interprétation par les différents médecins internes au laboratoire du sommeil, mais aussi externes, généralistes ou spécialistes, amenés à recevoir les comptes-rendus polysomnographiques et à adapter, éventuellement, le traitement pour un patient donné. Enfin, le patient lui-même peut vouloir recevoir un compte-rendu médical qu'il puisse comprendre facilement.

À titre de comparaison, le logiciel <sup>TM</sup>Somnologica propose de générer un compte-rendu polysomnographique de plusieurs pages représenté sur la figure 5.1. Ce compte-rendu de plusieurs pages est constitué de plusieurs sections affichant les résultats pour un événement donné ; les apnées, les désaturations, les micro-éveils sont traités indépendamment dans des parties séparées.

Par l'utilisation d'événements symboliques et d'un référentiel riche sémantiquement, la visualisation d'information s'inscrit, selon moi, parfaitement comme une méthode de fusion symbolique. Elle permet d'augmenter le niveau de connaissance des données.

FIGURE 5.1 – Compte-rendu polysomnographique du logiciel <sup>TM</sup>Somnologia

# 1 OptiSAS

## 1.1 Méthode

La démarche OptiSAS a pour objectif une représentation graphique des données polysomnographiques utilisant un référentiel pertinent afin de pouvoir en extraire une information riche sémantiquement.

Nous avons donc mis au point un référentiel dans cet esprit.

Ce référentiel a été construit en utilisant la méthode *Generalized CaseView method* (GCm) [42–45] qui est une méthode de visualisation de données permettant de mettre en évidence les groupes homogènes particulièrement clairement, et permet, en conséquence, de les identifier facilement et rapidement. Le référentiel est une carte pixelisée en 2D et la position de chaque individu se fait en fonction de 3 critères :

- le critère **binaire** qui permet de diviser l'espace en 2 ; on lui associe un paramètre bivalué. Chaque valeur est associée à un des côtés.
- le critère **nominal** qui permet de diviser l'espace en colonnes ; on lui associe un paramètre multivalué dont le nombre de valeurs est fini. À chaque valeur différente est associée une nouvelle colonne.
- le critère **ordinal** qui permet de savoir à quelle distance de l'axe médian le point se situera ; on lui associe un paramètre à valeurs dans un ensemble où il existe une relation d'ordre. De façon courante, on utilise un paramètre à valeurs numériques.

Chaque cellule correspond ainsi à un pixel auquel est affectée une couleur, dépendant de la valeur d'un dernier paramètre, selon une échelle de couleurs prédéfinie. Ce pixel est appelé *infoxel* [44].

La figure 5.2 illustre la méthode GCm et ses 3 paramètres de localisation spatiale.

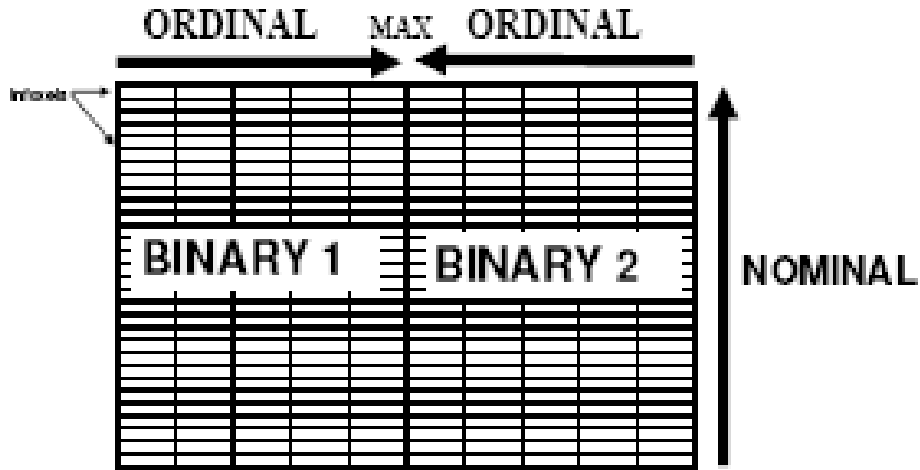


FIGURE 5.2 – Méthode GCm

Dans notre cas, l'application de la méthode GCm a conduit à la définition des critères suivants :

- Le critère binaire est la position de sommeil. Pour obtenir un critère binaire, nous avons mis sur la gauche de l'axe médian, tout ce qui était lié à la position *décubitus dorsal* (*supine position*) et à droite tout ce qui était lié aux 3 autres positions (*décubitus ventral*, *décubitus latéral droit* et *décubitus latéral gauche*).
- Le critère nominal est l'ensemble des événements que l'expert médical souhaite surveiller. Classiquement, sont surveillés : les apnées et hypopnées confondues, les apnées (tous types confondus), les apnées obstructives, les apnées centrales, les apnées mixtes, les hypopnées, les désaturations, les micro-éveils et éventuellement les ronflements.
- Le critère ordinal est le stade de sommeil : sommeil paradoxal (proche de l'axe), sommeil lent léger, sommeil lent profond (aux extrémités).

La définition de ce référentiel est issue des travaux de Slawomir PIETRASZ, qui, en collaboration avec Carole PHILIPPE, lors d'un stage ingénieur, a validé une idée du Pierre LÉVY [65]. J'ai poursuivi ensuite ces travaux lors de l'un de mes stages ingénieur [74].

L'information renfermée par l'infoxel est l'index de l'événement correspondant au critère nominal ; l'échelle de couleur utilisée est basée sur les 4 classes de gravité de Syndrome d'Apnées du Sommeil reconnues dans la communauté des médecins du sommeil (absence de SAS, SAS léger, SAS modéré, SAS sévère) :

- $index < 5$  : blanc ;
- $5 \leq index < 15$  : jaune ;
- $15 \leq index < 30$  : orange ;
- $index > 30$  : rouge.

Afin de faciliter l'interprétation, la largeur des pixels a été ajustée proportionnellement à la durée de sommeil dans chaque contexte de sommeil (stade de sommeil / position).

Cette méthode a été publiée dans [77] et présentée oralement lors de la *20th Conference of the European Federation for Medical Informatics (MIE 2008)* qui s'est tenue à Göteborg, en Suède, du 25 au 28 Mai 2008.

## 1.2 Résultats

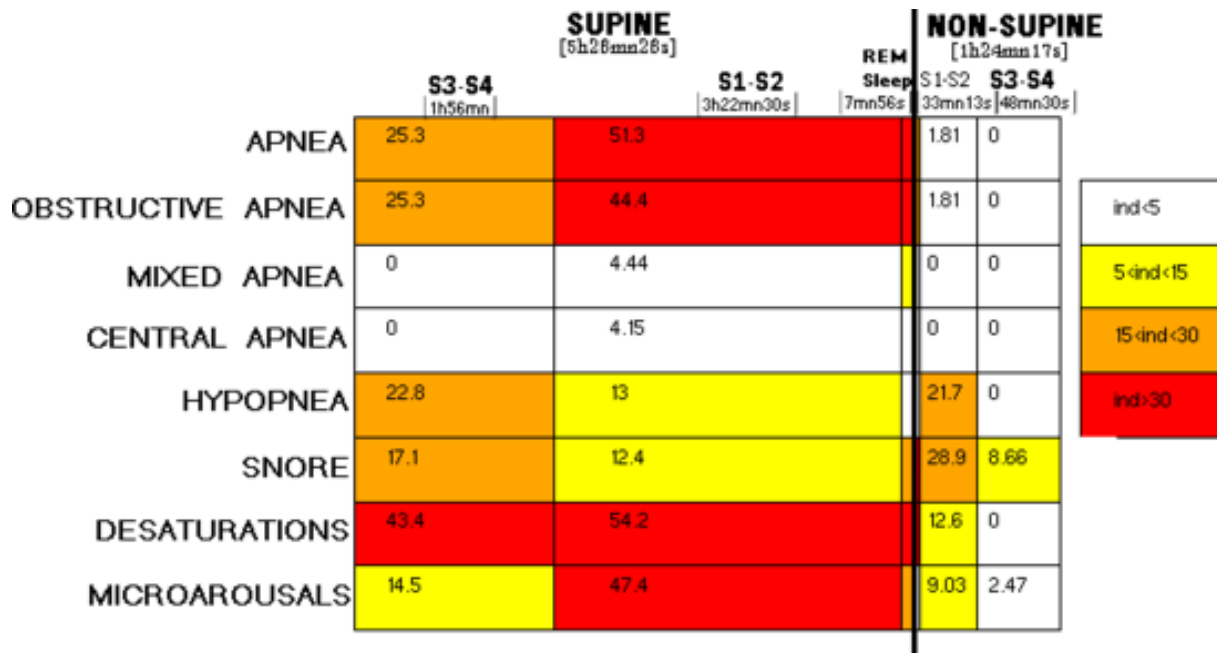


FIGURE 5.3 – Exemple de graphique OptiSAS

La figure 5.3 représente le graphique OptiSAS de l'examen polysomnographique d'un patient.

On peut y voir que le patient a dormi 5h26mn26s sur le dos et 1h24mn17s dans les autres positions. Il a très peu dormi en sommeil paradoxal (7mn56s). Les temps de sommeil dans les autres stades de sommeil sont conformes à la normale.

On peut y voir que ce patient fait de nombreuses apnées qui sont toutes quasiment exclusivement de type obstructif. Il y a très peu d'apnées centrales et d'apnées mixtes. Le patient fait également des hypopnées, mais celles-ci sont quasiment 3 fois moins fréquentes en sommeil lent léger dans la position décubitus dorsal, contexte dans lequel le patient a passé la moitié de son temps de sommeil.

Ces apnées causent à la fois des désaturations et des micro-éveils; on peut le voir à leurs index respectifs élevés.

On peut également clairement voir la prédominance des événements en position décubitus dorsal. Cela indique un SAS positionnel.

Par contre, il ne semble pas y avoir de forte prédominance d'un stade de sommeil par rapport à un autre, les différences d'index entre le sommeil lent léger et le sommeil lent profond étant habituelles.

En conclusion, ce patient souffre d'un Syndrome d'Apnées Obstructives du Sommeil sévère avec forte prédominance en position dorsale, associé à des index élevés de désaturations et de micro-éveils.

### 1.3 Évaluation

Nous avons soumis la méthode OptiSAS à deux évaluations. La première évaluation a consisté à donner à Carole PHILIPPE, médecin responsable de l'unité d'explorations fonctionnelles du sommeil de l'Hôpital Tenon, les graphiques OptiSAS de 13 patients souffrant de Syndrome d'Apnées du sommeil. Pour chacun d'entre eux, elle devait dire si le graphique lui permettait de découvrir un lien entre le SAS et une position et/ou un stade de sommeil. Pour 7 patients (53,8%), la dépendance à un stade de sommeil a été découverte ; pour 4 patients (30,8%), la dépendance à une position a été mise en évidence. Globalement, le graphique OptiSAS a permis d'obtenir de nouvelles informations pour 8 patients, soit 62% des patients.

La deuxième évaluation a été effectuée avec 4 médecins attachés à l'unité d'explorations fonctionnelles du sommeil de l'Hôpital Tenon ; certains étaient plus spécialistes en neurologie, d'autres en pneumologie. Chacun a reçu 20 graphiques OptiSAS. Le groupe de 20 patients correspondant était constitué de 14 Hommes et 6 Femmes. L'âge moyen était de  $62,05 \pm 17,32$  ans ; l'IAH moyen était de  $31 \pm 14,95$ . L'Indice de Masse Corporelle moyen ( $\frac{\text{masse}}{\text{taille}^2}$ ) était de  $25,87 \pm 3,05 \text{ kg/m}^2$ . Pour chaque cas, le médecin devait répondre à 3 questions :

1. « Quel est le type de SAS ? » ;
2. « Quelle est la sévérité du SAS ? » ;
3. « Existe-t-il une dépendance au stade de sommeil ? » ;
4. « Existe-t-il une dépendance à la position ? ».

Nous avons mesuré le coefficient kappa de concordance de chaque expert avec l'expert de référence (qui avait à sa disposition l'ensemble des informations et résultats concernant chaque patient) pour chacune des 4 questions. Les résultats sont donnés dans le tableau 5.1.

TABLEAU 5.1 – Coefficient Kappa de concordance des experts avec l'expert de référence pour l'évaluation du graphique OptiSAS

Expert ID	Type de SAS	Sévérité du SAS	Dépendance à la position	Dépendance au stade de sommeil
1	0,53	0,38	0,38	0,44
2	0,33	0,54	0,34	-0,07
3	0,22	0,35	0,5	0,09
4	0,49	0,29	0,24	-0,15

Plus en détails, nous avons constaté que la concordance sur le critère « Type de SAS » est modérée pour 50% des experts et pauvre pour les autres 50%.

Pour les critères « Sévérité du SAS » et « Dépendance à la position », la concordance est modérée pour 25% des experts et pauvre pour les 75% restants.

Pour le critère « Dépendance au stade de sommeil », la concordance est modérée pour 25% des experts, quasi nulle pour 25% et négative (quasi nulle) pour les 50% restants.

## 1.4 Discussion

Nous constatons donc que par une analyse unique du graphique OptiSAS, nous avons la meilleure concordance pour le type de SAS, en particulier dans le cas des SAS obstructifs. Les concordances faibles pour les autres critères ont permis de mettre en évidence que chaque médecin a une sensibilité différente à la sévérité ou à la dépendance à la position ou à un stade de sommeil, chacun utilise des seuils différents.

Ces conclusions justifient donc le recours à un support standardisé de représentation et communication des résultats. Le graphique OptiSAS semble aller dans ce sens. Assez petit, rapide et simple à interpréter, il contient de nombreuses informations qui permettent un premier niveau de diagnostic. Les différences constatées lors de l'évaluation entre les différents experts doivent nous alerter sur la nécessité que les experts s'habituent à travailler avec cette représentation. Pour faciliter la lecture, il faut peut-être ajouter les graphiques OptiSAS de profils types élaborés par une équipe d'experts de médecine du sommeil.

Ces résultats ont été publiés [76] et présentés oralement à la *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society* à Minneapolis, dans l'État de Minnesota aux États-Unis du 3 au 6 Septembre 2009.

## 2 SAS-Map

La démarche SAS-Map est dans la continuité de la démarche OptiSAS. Il s'agit ici d'exploiter les résultats de la fusion symbolique pour leur représentation graphique et d'adapter la disposition des informations à la lecture de l'expert médical. Comme dans la démarche OptiSAS, la *SAS-Map* doit permettre une interprétation rapide et précise du contenu et donc utiliser un référentiel riche sémantiquement.

## 2.1 Méthode

La *SAS-Map* est divisée en 3 sections :

1. La première section donne des informations sur les temps de sommeil, par contexte de sommeil (binôme stade de sommeil / position).

Tout d'abord, le temps de sommeil total est écrit.

Ensuite, il est indiqué, par un histogramme horizontal, pour chaque stade de sommeil, le temps de sommeil et sa répartition par position.

Enfin, il est indiqué, par un autre histogramme horizontal, pour chaque position, le temps de sommeil et sa répartition par stade de sommeil.

2. La seconde section donne des informations sur les événements respiratoires. Elle est constituée de trois parties. La première donne des informations sur l'IAH ; la deuxième précise les index pour les apnées et les hypopnées ; la troisième précise la répartition pour chaque type d'apnées. Chaque graphique est construit selon le même référentiel : l'espace est divisé en quatre cadrans par un axe horizontal et un axe vertical. Chaque cadran correspond à une position. En partant de l'axe médian vertical, les segments sont ordonnés par stade de sommeil : REM, N1, N2 puis N3. La largeur de chaque segment est proportionnelle au temps de sommeil passé dans le contexte associé. Chaque segment peut alors être associé à un infoxel du graphique OptiSAS et renferme comme information un index. Pour faciliter l'interprétation de chaque segment, une couleur de fond lui est associée, respectant une échelle de couleur continue, à la différence d'OptiSAS qui avait une échelle de couleur discrète par classe de sévérité.

Un graphique est donné pour l'IAH avec une ligne par cadran. Un graphique est également donné pour les événements respiratoires avec deux lignes par cadran : une pour les apnées et une pour les hypopnées. Enfin, 4 graphiques sont donnés pour le type d'apnées, un pour chaque type (obstructives, centrales, mixtes et indécidables)

3. La dernière section donne des informations sur les conséquences physiologiques entraînées par les événements respiratoires. Les apnées sont encore triées par type, avec un graphique par type. L'information renfermée par chaque segment, c'est-à-dire ici pour le contexte (type d'apnées / position / stade de sommeil) est un pourcentage. Ainsi, on peut connaître le pourcentage d'apnées obstructives survenues en stade N2 et en position décubitus dorsal ayant entraîné une désaturation ou un micro-éveil.

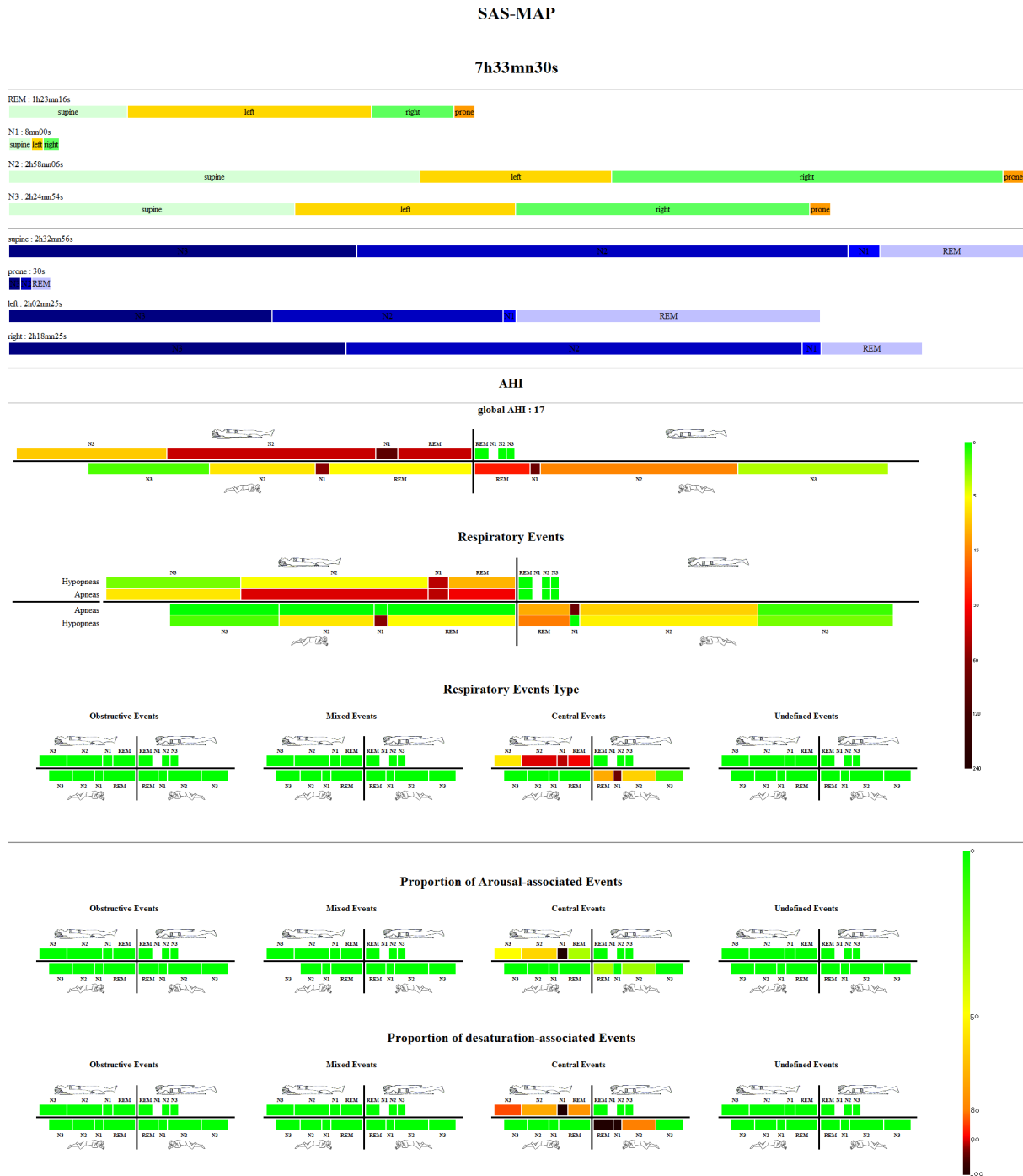


FIGURE 5.4 – Exemple de SAS-Map



## 2.2 Résultats

La figure 5.4 est un exemple de SAS-Map, construit à partir de l'analyse automatique de l'enregistrement d'un patient.

On peut y voir que le patient a dormi 7h33mn30s dont 8mn (1,9%) passées en stade N1, 2h58mn06s (43,0%) passées en stade N2, 2h24mn54s (35,0%) passées en stade N3 et 1h23mn16s (20,1%) passées en sommeil paradoxal. C'est une répartition à peu près normale, malgré un temps passé en sommeil profond (N3) supérieur à la normale, qui peut s'expliquer par une dette de sommeil. Nous constatons aussi que ce patient a dormi dans toutes les positions, et ceci pour chaque stade de sommeil. On peut voir que le patient a également passé une durée équivalente dans chaque position, hormis la position décubitus ventral, ce qui est un résultat habituel, surtout avec le polysomnographie.

L'IAH moyen est de 17, ce qui est une indication d'un SAS modéré. On peut voir que les événements surviennent plus fréquemment en position décubitus dorsal, moins fréquemment en position décubitus latéral gauche, mais quasiment jamais en positions décubitus ventral et décubitus latéral droit. Cela indique un SAS positionnel favorisé par la position décubitus dorsal. De plus, les IAH en stade N1 sont très élevés, ce qui n'est pas très significatif compte tenu de la faible durée du stade N1. Les IAH pour les stades de sommeil N2 et REM sont comparables ; la valeur plus faible de l'IAH en stade N3 est habituelle.

Nous constatons aussi qu'en position décubitus dorsal, les événements majoritaires sont des apnées ; en position décubitus latéral gauche, il y a autant d'apnées que d'hypopnées.

La dernière partie indique que les apnées sont exclusivement de type central.

Enfin, on peut voir que les apnées sont assez rarement suivies de micro-éveils ou de désaturations.

En conclusion, ce patient a un Syndrome d'Apnées du Sommeil modéré de type central, favorisé par la position *décubitus dorsal*. Les événements majoritaires sont des apnées qui entraînent peu de désaturations ou de micro-éveils.

La méthode SAS-Map a fait l'objet d'un poster présenté lors de la *XXIII International Conference of the European Federation for Medical Informatics* (MIE 2011) qui a eu lieu à Oslo en Norvège du 28 au 31 Août 2011.

## 2.3 Discussion

Nous n'avons encore procédé à aucune évaluation de la SAS-Map. Cependant, il semble que son interprétation satisfasse les exigences de simplicité et de richesse d'interprétation que nous nous étions fixées au départ. Nous profitons aussi pleinement de la fusion symbolique en conjuguant le référentiel (position / stade de sommeil / type d'apnées) aux différents de niveaux de lecture avec une précision d'informations qui augmente quand on lit de haut en bas (temps de sommeil -> IAH -> événements respiratoires séparés -> type d'apnées -> conséquences physiologiques).

L'échelle de couleur continue permet de se rapprocher de la réalité.

## 3 Conclusion

Dans l'objectif de standardiser la représentation des résultats de l'analyse d'un examen polysomnographique en utilisant un référentiel qui soit à la fois riche en informations mais aussi simple et rapide à interpréter, pour qu'il satisfasse les besoins, différents, de l'équipe de médecins du laboratoire du sommeil, les médecins partenaires et les patients, le graphique OptiSAS et la SAS-Map apportent des éléments de réponse. Ils permettent chacun de bénéficier des résultats de la fusion symbolique en incluant les contextes de survenue des événements respiratoires et, dans le cas de la SAS-Map, leurs conséquences physiologiques.

L'évaluation d'OptiSAS a montré que, malgré tout, chaque expert a une interprétation différente des mêmes résultats. Standardiser la représentation n'est donc pas suffisante. Il s'agit donc de poursuivre ces travaux, en ajoutant aux graphiques OptiSAS des graphiques de référence correspondant à des profils types de malades qu'il suffirait ensuite, de rapprocher visuellement. Ceci pourrait se faire automatiquement par le calcul d'une distance entre deux graphiques OptiSAS et ainsi, guider un peu plus dans la lecture. Cela suppose que l'interprétation des graphiques de profils types aura un meilleur consensus.

Enfin, il faut évaluer la SAS-Map pour orienter des recherches futures.

# Conclusion

## Bilan

Le travail de recherche présenté dans ce mémoire concerne l'utilisation de la fusion symbolique à la résolution des problèmes biomédicaux. Le secteur biomédical est un domaine spécifique où l'erreur doit être absolument évitée et c'est pourquoi les outils utilisés servent de support au diagnostic et non d'outil de diagnostic. L'expert médical garde toujours un contrôle et une validation finale des résultats. Les moyens techniques utilisés pour le diagnostic sont de plus en plus complexes et il est courant de recourir à plusieurs examens pour effectuer le diagnostic approfondi d'une pathologie. Cela pose des difficultés car le diagnostic nécessite la fusion des informations issues des résultats des différents examens, et les médecins, même spécialistes, n'en maîtrisent pas toujours toutes les subtilités. La décision est donc collégiale et profite des résultats des différents examens, du contexte clinique du patient et des connaissances de chaque expert dans sa spécialité.

Dans son rapport sur le thème du sommeil remis au Ministre de la Santé en Décembre 2006, le Dr J.P. GIORDANELLA constate que la médecine du sommeil, « discipline jeune, [...] s'est largement développée dans la dernière décennie du fait de l'émergence du Syndrome d'Apnées du Sommeil comme problématique de santé publique avec ses conséquences sur nombre de spécialités d'organes ou de systèmes ». En effet, le Syndrome d'Apnées du sommeil a une prévalence estimée à 2% de la population des Femmes d'âge moyen et 4% de la population des Hommes d'âge moyen. Ses conséquences incluent une somnolence diurne excessive, démontrée comme étant l'une des causes majeures de survenue d'accident de la route ou du travail, mais aussi un surcroît de morbidité et de mortalité cardio et cérébro-vasculaires. L'examen médical de référence pour effectuer son diagnostic s'appelle la polysomnographie et consiste en l'enregistrement simultané, pendant toute une nuit, de plusieurs paramètres physiologiques incluant, entre autres, l'activité cérébrale, le flux respiratoire et l'oxymétrie. Selon les recommandations officielles en France, l'interprétation des signaux doit être *a minima* validée par un expert médical. L'interprétation des signaux nécessite souvent l'intervention de plusieurs experts du sommeil, certains avec une spécialité à dominante en pneumologie, d'autres plus spécialistes des aspects neurologiques. Il s'agit d'une tâche chronophage et fastidieuse, dont le résultat présente quelques divergences inter et intra-scoreurs, malgré la standardisation en 2007 des méthodes d'interprétation par l'American Academy of Sleep Medicine (AASM). Des chercheurs ont démontré que ces divergences pouvaient être atténuées par l'utilisation d'un dispositif de support au diagnostic

Le support au diagnostic du Syndrome d'Apnées du Sommeil est donc un contexte de travail favorable à l'utilisation de la fusion, et plus particulièrement de la fusion

symbolique. En effet, la fusion symbolique utilise l'information, représentée sous un formalisme symbolique, issue de plusieurs sources et la fusionne, conformément à des règles et des stratégies de fusion, pour obtenir une connaissance. Cette démarche est très proche de ce qui est déjà effectué par les médecins du sommeil pour effectuer le diagnostic. Les troubles respiratoires pathologiques symptomatiques du Syndrome d'Apnées du Sommeil ont un contexte de survenue et des conséquences physiologiques qui doivent être analysés précisément pour effectuer un diagnostic approfondi et faciliter la décision de traitement. C'est ce que permet la fusion symbolique.

La fusion symbolique nécessite une étape préalable, de recueil des événements symboliques à fusionner. Dans le cas du Syndrome d'Apnées du Sommeil, ils sont de deux types : il existe des événements courts pathologiques et des événements contextuels, c'est-à-dire des conditions environnementales aux événements courts. Nous avons démontré que nous pouvions utiliser des méthodes efficaces pour le recueil de ces événements. Ces algorithmes sont basés sur une première étape utilisant des algorithmes de traitement du signal.

Dans un premier temps, nous avons présenté deux méthodes de reconnaissance automatique des stades de sommeil. La première utilise un arbre binaire de décision sur des paramètres fréquentiels extraits des signaux neurologiques enregistrés. La seconde utilise la fusion symbolique. Ces deux méthodes s'avèrent efficaces et permettent d'obtenir des résultats à peu près équivalents à ceux obtenus par un autre scoreur humain. Il est intéressant de constater que pour améliorer les résultats, il serait nécessaire dans les deux cas de concevoir de nouveaux paramètres, mais que ceux-ci pourraient être adjoints directement au système existant par la conjonction de nouvelles règles dans le cas de la fusion symbolique, alors que l'arbre binaire de décision nécessite une refonte totale de l'arbre pour la prise en compte de ces nouveaux paramètres.

Dans un second temps, nous avons procédé à la détection automatique des événements respiratoires et à leur classification en apnées - obstructives, centrales, mixtes ou indécidables - et en hypopnées. La comparaison des résultats montre que la reconnaissance des événements respiratoires se fait correctement, malgré parfois un étiquetage différent de l'événement. La reconnaissance du type d'apnées est également correcte pour les apnées obstructives et les apnées centrales, mais semble mauvaise pour identifier les apnées mixtes. Ces différences seront vues avec un expert médical en vue d'une amélioration. La reconnaissance automatique des désaturations, plus simple, donne d'excellents résultats.

Nous avons ensuite démontré que la fusion de tous ces événements était efficace car elle permettait d'obtenir des résultats équivalents à ceux obtenus par un expert médical. La fusion symbolique va même plus loin en mettant en évidence des liens entre les événements qui aujourd'hui ne sont pas utilisés, ce qui pose des difficultés pour une validation médicale ; la validation algorithmique, elle, est effectuée. Ceci pose aussi la question de la restitution des résultats. Pour y répondre, nous avons proposé, dans le dernier chapitre, deux méthodes originales de visualisation des résultats, selon un référentiel permettant à l'œil expérimenté une intégration de toute l'information sémantique renfermée dans les indicateurs présentés. Une validation médicale reste nécessaire après une appropriation de ces référentiels par l'expert.

Sur le plan scientifique, ce doctorat a été très riche. L'aspect pluridisciplinaire m'a permis de travailler avec des experts de domaines différents, issus de plusieurs laboratoires en France, mais aussi à l'étranger, au travers d'une collaboration avec l'Université de Ballarat en Australie. Ces aspects collaboratifs participent à la richesse humaine de cette thèse.

## **perspectives**

Il existe de nombreuses pistes de poursuite des travaux présentés dans ce mémoire.

Dans le but d'une utilisation routinière, le premier objectif est l'évaluation d'OptiSAS et de SAS-Map sur des nouveaux patients pour connaître les bénéfices et les inconvénients de la fusion symbolique et des 2 référentiels présentés.

Il serait également intéressant d'utiliser l'analyse automatique comme pré-analyse et de demander, ensuite, à deux experts de procéder à une interprétation en appliquant, éventuellement, des modifications aux résultats de la pré-analyse. L'évaluation consistera alors à mesurer le taux de modifications ainsi que le taux d'accord entre les interprétations finales des deux experts.

Il semble aussi nécessaire d'ajouter un module de pré-traitement des courbes enregistrées afin de voir s'il est possible d'en réduire le bruit, ou, à défaut, de déclencher une alerte pour informer de l'incapacité de la méthode à analyser, en toute fiabilité, cet enregistrement polysomnographique.

Enfin, afin d'apporter de nouvelles connaissances aux résultats, il faudrait créer de nouveaux algorithmes de reconnaissance automatique des micro-éveils, des mouvements de jambes ou des événements cardiaques. De même, en reconnaissant automatiquement les motifs très courts spécifiques du sommeil lent léger comme les complexes K ou les pointes vertex, il serait alors possible d'améliorer notre méthode utilisant la fusion, par l'ajout de nouveaux paramètres.

Il existe enfin des voies d'amélioration des algorithmes de détection et classification des événements respiratoires. Profitant de l'analyse approfondie, avec un expert, de chacune des erreurs et des différences de classification, il sera alors possible d'apporter des modifications et des corrections pertinentes, en définissant par exemple de nouveaux paramètres pertinents et en ajoutant de nouvelles règles.



# Bibliographie

- [1] Pierre-Antoine ABSIL : Analyse non linéaire de signaux cardiologiques en vue d'applications cliniques. Mémoire de D.E.A., Université de Liège (Belgique), 1998.
- [2] Pierre-Antoine ABSIL, Rodolphe SEPULCHRE, Ali BILGE et Robert POIRRIER : Sleep apnea syndrome diagnosis by analysis of the cardiac interbeat RR time series. In J.A.K. SUYKENS et J. VANDEWALLE, éditeurs : *Proceedings of the International Workshop on Advanced Black-Box Techniques for Nonlinear Modeling*, pages 133–137. Katholieke Universiteit Leuven, Belgium, July 8-10 1998.
- [3] Tarik AL-ANI, Yskandar HAMAM, Daniel NOVAK, Pala Tamara POZZO MENDOZA, Lenka LHOTSKA, Frédéric LOFASO, Daniel ISABEY et Redouane FODIL : Noninvasive automatic sleep apnea classification system. In *Proceedings of the Conference on Modeling and Simulation in Biology, Medicine and Biomedical Engineering*, pages 45–55, 2005.
- [4] Tarik AL-ANI, Radek KAZBUNDA et Daniel NOVÁK : Automatic sleep scoring based only on electrocardiogram records. In *Proceedings of the 6<sup>th</sup> EUROSIM Congress on Modeling and Simulation [9-13 September, 2007, Ljubljana, Slovenia]*, 2007.
- [5] Peter ANDERER, Georg GRUBER, Silvia PARAPATICS, Michael WOERTZ, Tatiana MIAZHYNKAIA, Gerhard KLOSCH, Bernd SALETU, Josef ZEITLHOFFER, Manuel J. BARBANOJ, Heidi DANKER-HOPFE, Sari-Leena L. HIMANEN, Bob KEMP, Thomas PENZEL, Michael GROZINGER, Dieter KUNZ, Peter RAPPELSBERGER, Alois SCHLOGL et Georg DORFFNER : An e-health solution for automatic sleep classification according to Rechtschaffen and Kales : validation study of the somnolyzer 24 x 7. *Neuropsychobiology*, 51(3):115–133, 2005.
- [6] Michael ARZT, Terry YOUNG, Laurel FINN, James B. SKATRUD et T. Douglas BRADLEY : Association of sleep-disordered breathing and the occurrence of stroke. *Am J Respir Crit Care Med*, 172(11):1447–1451, 2005.
- [7] <http://www.ballarat.edu.au/ard/itms/CIAO/ganso/>.
- [8] Isaac N. BANKMAN, Vincent G. SIGILLITO, Robert A. WISE et Philip L SMITH : Feature-based detection of the k-complex wave in the human electroencephalogram using neural networks. *IEEE Trans Biomed Eng*, 39(12):1305–10, 1992.
- [9] Gleb BELIAKOV et Julien UGON : Implementation of novel methods of global and nonsmooth optimization : Ganso programming library. *Optimization*, 56(5-6):543–546, October-December 2007.
- [10] Christian BERTHOMIER, Xavier DROUOT, Maria HENMAN-STOICA, Pierre BERTHOMIER, Jacques PRADO, Djibril BOLAR-THIRE, Odile BENOIT, Jérémie MATTOUT et Marie-Pia D'ORTHO : Automatic analysis of single-channel sleep eeg : Validation in healthy individuals. *Sleep*, 30(11):1587–1595, Nov 2007.
- [11] Christian BERTHOMIER, Jacques PRADO et Odile BENOIT : Automatic sleep eeg analysis using filter banks. *Biomed Sci Instrum*, 35:241–6, 1999.

- [12] Pierre-Olivier BRIDEVAUX, Jean-William FITTING, Jean-Marc FELLRATH et John-David AUBERT : Inter-observer agreement on apnoea hypopnoea index using portable monitoring of respiratory parameters. *Swiss Med Wkly*, 137(43-44):602–607, 2007.
- [13] Mariano CABRERO-CANOSA, Elena HERNANDEZ-PEREIRA et Vicente MORET-BONILLO : Intelligent diagnosis of sleep apnea syndrome. *IEEE Engineering in medicine and biology Magazine*, March/April 2004.
- [14] Nitesh V. CHAWLA, Kevin W. BOWYER, Lawrence O. HALL et W. Philip KEGELMEYER : Smote : Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.
- [15] Jacob COHEN : A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37, 1960.
- [16] Nancy A COLLOP : Scoring variability between polysomnography technologists in different sleep laboratories. *Sleep Med*, 3(1):43–47, 2002.
- [17] Heidi DANKER-HOPFE, Peter ANDERER, Josef ZEITLHOFER, Marion BOECK, Hans DORN, Georg GRUBER, Esther HELLER, Erna LORETZ, Doris MOSER, Silvia PARAPATICS, Bernd SALETU, Andrea SCHMIDT et Georg DORFFNER : Interrater reliability for sleep scoring according to the rechtschaffen & kales and the new aasm standard. *Journal of sleep research*, 18(1):74–84, mar 2009.
- [18] Heidi DANKER-HOPFE, Dieter KUNZ, Georg GRUBER, Gerhard KLÖSCH, José L LORENZO, Sari-Leena HIMANEN, Bob KEMP, Thomas PENZEL, Joachim RÖSCHKE, Hans DORN, Alois SCHLÖGL, E TRENKER et Georg DORFFNER : Interrater reliability between scorers from eight european sleep laboratories in subjects with different sleep disorders. *J Sleep Res*, 13(1):63–9, 2004.
- [19] Belur V DASARATHY : Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proceedings of the IEEE*, 85(1):24–38, 1997.
- [20] Michael K DRINNAN, Alan MURRAY, Clive K GRIFFITHS et G JOGN GIBSON : Interobserver variability in recognizing arousal in respiratory sleep disorders. *Am. J. Respir. Crit. Care Med.*, 158(2):358–362, 1998.
- [21] Ángel FERNÁNDEZ-LEAL et Vicente MORET-BONILLO : Temporal knowledge in the intelligent diagnosis of sleep apnea syndrome : a new approach. *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 1039–1042, aug 2007.
- [22] R FERRI, P FERRI, R M COLOGNOLA, M A PETRELLA, S A MUSUMECI et P BERGONZI : Comparison between the results of an automatic and a visual scoring of sleep eeg recordings. *Sleep*, 12(4):354–62, 1989.
- [23] Luay A. FRAIWAN, Natheer Y. KHASWANEG et Khaldon Y. LWEESEY : Automatic sleep stage scoring with wavelet packets based on single eeg recording. *World Academy of Science Engineering and Technology*, 54:485–488, 2009.
- [24] Ana FRED, Joaquim FILIPE, Markku PARTINEN et Teresa PAIVA : Psg-expert. an expert system for the diagnosis of sleep disorders. *Stud Health Technol Inform*, 78:127–47, 2000.
- [25] Charles F.P. GEORGE : Reduction in motor vehicle collisions following treatment of sleep apnoea with nasal cpap. *Thorax*, 56(7):508–512, jul 2001.
- [26] Jean-Pierre GIORDANELLA : Rapport sur le thème du sommeil. Rapport technique, Ministère de la Santé et des Solidarités, Dec 2006.



- 
- [27] Gabriela GUIMARÃES, Jörg-Hermann PETER, Thomas PENZEL et Alfred ULTSCH : A method for automated temporal knowledge acquisition applied to sleep-related breathing disorders. *Artificial Intelligence in Medicine*, 23(3):211–237, novembre 2001.
  - [28] S GÜNEŞ, K POLAT et Ş YOSUNKAYA : Efficient sleep stage recognition system based on eeg signal using k-means clustering based feature weighting. *Expert Syst. Appl.*, 37:7922–7928, Dec 2010.
  - [29] Bassel HABIB, Claire LAUDY et Jean-Gabriel GANASCIA : Using fusion to fill in the gaps in old scientific discoveries’ notebooks. In *Proceedings of IJCAI-09 Workshop on Graph Structures for Knowledge Representation and Reasoning (GKR 2009) of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2009)*, 2009.
  - [30] Masaaki HANAOKA, Masaki KOBAYASHI et Haruaki YAMAZAKI : Automatic sleep stage scoring based on waveform recognition method and decision-tree learning. *Systems and Computers in Japan*, 33(11):1–13, 2002.
  - [31] Conrad IBER, Sonia ANCOLI-ISRAEL, Andrew L. CHESSON et Stuart F. QUAN : *The AASM Manual for the Scoring of Sleep and Associated Events : Rules, Terminology and Technical Specifications*. The American Academy of Sleep Medicine (AASM), Westchester, Illinois, USA, 1<sup>st</sup> ed. édition, 2007.
  - [32] Nizar KERKENI, Frédéric ALEXANDRE, Mohamed Hédi BEDOUI, Laurent BOUGRAIN et Mohamed DOGUI : Neuronal spectral analysis of eeg and expert knowledge integration for automatic classification of sleep stages. *SEAS Transactions on information Science Applications*, 2(11):1854–1861, 2005. <http://www.wseas.org>.
  - [33] Ahsan H. KHANDOKER, Jayavardhana GUBBI et Marimuthu PALANISWAMI : Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings. *Trans. Info. Tech. Biomed.*, 13(6):1057–1067, nov 2009.
  - [34] Kenji KIRA et Larry A. RENDELL : A practical approach to feature selection. In *Proceedings of the ninth international workshop on Machine learning*, ML92, pages 249–256, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.
  - [35] J. Richard LANDIS et Gary G. KOCH : The measurement of observer agreement for categorical data. *Biometrics*, 33:159–174, mar 1977.
  - [36] Claire LAUDY et Jean-Gabriel GANASCIA : Information fusion in a tv recommendation system. In *Proceedings of the 11th International Conference on information Fusion*, FUSION 2008, 2008.
  - [37] Claire LAUDY et Jean-Gabriel GANASCIA : Introducing semantic knowledge in high-level fusion. In *Proceedings of the 28th IEEE conference on Military communications*, MILCOM’09, pages 953–959, Piscataway, NJ, USA, 2009. IEEE Press.
  - [38] Claire LAUDY et Jean-Gabriel GANASCIA : Using maximal join for information fusion. In *Proceedings of IJCAI-09 Workshop on Graph Structures for Knowledge Representation and Reasoning (GKR 2009) of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2009)*, 2009.
  - [39] Claire LAUDY, Jean-Gabriel GANASCIA et Célestin SEDOGBO : High-level fusion based on conceptual graphs. In *10th International Conference on Information Fusion*, pages 8–12, 2007.
  - [40] Claire LAUDY, Bassel HABIB et Jean-Gabriel GANASCIA : Fusion of claude bernard’s experiments for scientific discovery reasoning. In *Proceedings of the 17th International Conference on Conceptual Structures : Conceptual Structures : Leveraging Semantic Technologies*, ICCS ’09, pages 219–232, Berlin, Heidelberg, 2009. Springer-Verlag.

- [41] Adrien-Marie LEGENDRE : Appendice sur la méthode des moindres carrés. *In Nouvelles méthodes pour la détermination des orbites des comètes*, pages 72–80. Firmin-Didot, Paris, 1805.
- [42] Pierre P LÉVY : Le case view, une méthode de visualisation du case mix. *Journal d'économie médicale*, 20:118–127, March 2002.
- [43] Pierre P LÉVY : The case view, a generic method of visualization of the case mix. *International Journal of Medical Informatics*, 73(9-10):713–8, septembre 2004.
- [44] Pierre P. LÉVY : Pixelization paradigm : outline of a formal approach. *In Proceedings of the 1st first visual information expert conference on Pixelization paradigm, VIEW'06*, pages 3–11, Berlin, Heidelberg, 2007. Springer-Verlag.
- [45] Pierre P LÉVY, Laetitia DUCHÉ, Laszlo DARAGO, Yves DORLÉANS, Laurent TOUBIANA, Jean-François VIBERT et Antoine FLAHAULT : ICPCview : visualizing the International Classification of Primary Care. *Studies in health technology and informatics*, 116:623–8, 2005.
- [46] H MA, B HU, J JACKSON, M andd Yan et W ZHAO : A hybrid classification method using artificial neural network based decision tree for automatic sleep scoring. *World Academy of Science, Engineering and Technology*, 79:279–284, 2011.
- [47] J. Víctor MARCOS, Roberto HORNERO, Daniel ÁLVAREZ, Félix del CAMPO et Mateo ABOY : Automated detection of obstructive sleep apnoea syndrome from oxygen saturation recordings using linear discriminant analysis. *Med. Biol. Engineering and Computing*, 48(9):895–902, 2010.
- [48] Jose M. MARIN, Santiago J. CARRIZO, Eugenio VICENTE et Alvar G. N. AGUSTI : Long term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnea with or without treatment with continuous positive airway pressure : an observational study. *The Lancet*, 365(9464):1046–1053, Mar 19-25 2005.
- [49] Wadie MASHAL : Obstructive sleep apnea analysis based on heart rate variability. Mémoire de D.E.A., Electrical and Computer Engineering, Portland State University, 2004.
- [50] Walter T. McNICHOLAS : Diagnosis of obstructive sleep apnea in adults. *Proc Am Thorac Soc*, 5(2):154–160, 2008.
- [51] Medcare Flaga hf, Flaga hf. Medical Devices - Thorsgata 24 - IS-101 Reykjavík - Iceland. *EMBLA File Format Description*, Nov 1997.
- [52] Manuel MERINO, Octavio RIVERA, Isabel GOMEZ, Alberto MOLINA et Enrique DORRONZORO : A method of eog signal processing to detect the direction of eye movements. *In Proceedings of the 2010 First International Conference on Sensor Device Technologies and Applications, SENSORDEVICES '10*, pages 100–105, Washington, DC, USA, 2010. IEEE Computer Society.
- [53] Joseph E MIETUS, Chung-Kang PENG, Plamen Ch IVANOV et Ary Louis GOLDBERGER : Detection of obstructive sleep apnea from cardiac interbeat interval time series. *Computers in Cardiology*, 27:753–756, 2000.
- [54] Isabel MILHO et Ana FRED : A user-friendly development tool for medical diagnosis based on bayesian networks. *In Proceedings of the Second International Conference on Enterprise Information Systems*, pages 4–7, 2000.
- [55] Ahmad Taufiq MOHAMAD : Exploration des stades de sommeil par l'analyse fréquentielle du signal électro-encéphalographique. Mémoire de D.E.A., ESIEE, 2009.



- [70] John F. SOWA : *Conceptual Structures : Information Processing in Mind and Machine*. Addison-Wesley, 1984.
- [71] Nadezda SUKHORUKOVA, Andrew STRANIERI, Bahadorreza OFOGHI, Peter VAM-  
PLEW, M. SALEEM, Liping MA, Adrien UGON, Julien UGON, Nial MUECKE, Hélène  
AMIEL, Carole PHILIPPE, Ahmed BANI-MUSTAFA, Shamsul HUDA, Marcello BER-  
TOLI, Pierre LÉVY et Jean-Gabriel GANASCIA : Automatic sleep stage identification :  
difficulties and possible solutions. *In HIC2009*, 2009.
- [72] Masaaki SUZUKI, Hanako SAIGUSA, Shintaro CHIBA, Tomoko YAGI, Kana SHIBA-  
SAKI, Mineko HAYASHI, Michiko SUZUKI, Kiyoshi MORIYAMA et Kazuoki KODERA :  
Discrepancy in polysomnography scoring for a patient with obstructive sleep apnea  
hypopnea syndrome. *Tohoku J Exp Med*, 206(4):353–360, 2005.
- [73] Ana Claudia Tonelli de OLIVEIRA, Denis MARTINEZ, Luiz Felipe T VASCONCE-  
LOS, Sandro Cadaval GONÇALVES, Maria Carmo LENZ, Sandra Costa FUCHS, Mi-  
guel GUS, Erlon Oliveira ABREU-SILVA, Leila Beltrami MOREIRA et Flávio Danni  
FUCHS : Diagnosis of obstructive sleep apnea syndrome and its outcomes with home  
portable monitoring. *Chest*, 135(2):330–6, 2009.
- [74] Adrien UGON : Optisas - modélisation de données polysomnographiques et cliniques.  
Mémoire de D.E.A., Institut d’Informatique d’Entreprise, 2005.
- [75] Adrien UGON, Jean-Gabriel GANASCIA, Carole PHILIPPE, Hélène AMIEL et Pierre  
LÉVY : How to use symbolic fusion to support the sleep apnea syndrome diagnosis.  
*In Proceedings of the 13th conference on Artificial intelligence in medicine*, AIME’11,  
pages 45–54, Berlin, Heidelberg, 2011. Springer-Verlag.
- [76] Adrien UGON, Carole PHILIPPE, Jean-Gabriel GANASCIA, Dominique RAKOTONA-  
NAHARY, Hélène AMIEL, Jean-Yves BOIRE et Pierre LÉVY : Evaluating optisas, a  
visual method to analyse sleep apnea syndromes. *In EMBC2009*, septembre 2009.
- [77] Adrien UGON, Carole PHILIPPE, Slawomir PIETRASZ, Jean-Gabriel GANASCIA et  
Pierre P. LÉVY : Optisas a new method to analyse patients with sleep apnea syn-  
drome. *In Stig Kjær ANDERSEN, Gunnar O. KLEIN, Stefan SCHULZ et Jos AARTS*,  
éditeurs : *MIE*, volume 136 de *Studies in Health Technology and Informatics*, pages  
547–552. IOS Press, 2008.
- [78] J. C. VÁZQUEZ, W. H. TSAI, W. W. FLEMONS, A. MASUDA, R. BRANT, E. HAJDUK,  
W. A. WHITELAW et J. E. REMMERS : Automated analysis of digital oximetry in  
the diagnosis of obstructive sleep apnoea. *Thorax*, 55(4):302–307, apr 2000.
- [79] Huy Quan VU, Gang LI, Nadezda S. SUKHORUKOVA, Gliel BELIAKOV, Simon LIU,  
Carole PHILIPPE, Hélène AMIEL et Adrien UGON : K-complex detection using a  
hybrid-synergic machine learning method. *IEEE Transactions on Systems, Man,  
and Cybernetics, Part C : Applications and Reviews*, PP(99):1 –13, 2012.
- [80] T. YOUNG, M. PALTA, J. DEMPSEY, J. SKATRUD, S. WEBER et S. BADR : The oc-  
currence of sleep-disordered breathing among middle-aged adults. *The New England  
journal of medicine*, 328(17):1230–1235, avril 1993.
- [81] Terry YOUNG, Paul E. PEPPARD et Daniel J. GOTTLIEB : Epidemiology of obstruc-  
tive sleep apnea : a population health perspective. *American Journal of Respiratory  
and Critical Care Medicine*, 165(9):1217–1239, 2002.
- [82] Zhi-hua ZHOU et Min-ling ZHANG : Neural networks for multi-instance learning. *In  
Proceedings of the International Conference on Intelligent Information Technology*,  
2002.



---

## RÉSUMÉ

Au cours des dernières décennies, la médecine a nécessité des examens de plus en plus complexes pour effectuer un diagnostic et orienter vers un traitement. Il est même courant de pratiquer plusieurs examens dans des spécialités médicales différentes afin d'étudier une pathologie par des approches multiples et ainsi mieux la connaître. Cela pose des difficultés d'interprétation car les données sont parfois hétérogènes mais surtout souvent très pointues et leur traitement requiert une expertise du domaine.

Dans ce contexte, la fusion symbolique constitue une solution possible. En effet, elle a prouvé son efficacité à traiter des problèmes sur des niveaux d'abstraction aussi bien faibles qu'élevés et à élaborer une connaissance de haut niveau.

Cette thèse démontre l'efficacité de la fusion symbolique appliquée au traitement des données polysomnographiques pour l'élaboration d'un outil de support au diagnostic du Syndrome d'Apnées du Sommeil. Pour être diagnostiqué, ce trouble du sommeil nécessite une polysomnographie. Cet examen médical consiste en l'enregistrement simultané de divers paramètres physiologiques durant toute une nuit. Son interprétation nécessite l'annotation des courbes enregistrées par une analyse visuelle effectuée par un médecin spécialiste du sommeil, ce qui est une tâche chronophage et fastidieuse dont les résultats peuvent présenter quelques divergences d'un expert à l'autre. Le recours à un outil de support au diagnostic fiable permet d'augmenter le consensus. Cette thèse développe les étapes d'élaboration d'un tel outil.

---

## TITLE

### Fusion symbolique et données polysomnographiques

---

## ABSTRACT

In recent decades, medical examinations required to diagnose and guide to treatment became more and more complex. It is even a current practice to use several examinations in different medical specialties to study a disease through multiple approaches so as to describe it more deeply.

The interpretation is difficult because the data is both heterogeneous and also very specific, with skilled domain of knowledge required to analyse it.

In this context, symbolic fusion appears to be a possible solution. Indeed, it was proved to be very effective in treating problems with low or high levels of abstraction of information to develop a high level knowledge.

This thesis demonstrates the effectiveness of symbolic fusion applied to the treatment of polysomnographic data for the development of an assisted diagnosis tool of Sleep Apnea Syndrome. Proper diagnosis of this sleep disorder requires a polysomnography. This medical examination consists of simultaneously recording of various physiological parameters during a night. Visual interpretation is tedious and time consuming and there commonly is some disagreement between scorers. The use of a reliable support-to-diagnosis tool increases the consensus. This thesis develops stages of the development of such a tool.

---

ISBN :